

SOMMAIRE

Introduction	1
Partie 1 - Contexte et objectifs	2
1. Analyse de l'existant	2
1.1. Etat des connaissances sur la maladie et le vecteur	2
1.2. Modèles existants de répartition du vecteur	2
2. Analyse des besoins	3
3. Objectifs	3
Partie 2 - Approche opérationnelle	4
1. Organisation des données épidémiologiques	4
1.1. Géoréférencement des élevages ovins et bovins	4
1.2. Recueil des caractéristiques sur les élevages	4
1.3. Recueil de l'information sur l'état sanitaire des élevages	5
2. Extraction des paramètres environnementaux	5
2.1. Choix des paramètres environnementaux	5
2.2. Calcul d'un indice de végétation	6
2.3. Réalisation d'une carte de végétation à partir de l'image SPOT	6
2.3.1. Acquisition des relevés de végétation	7
2.3.2. Classifications	8
2.4. Extraction des paramètres environnementaux liés à la topographie	11
3. Analyse	11
3.1. Intégration des données dans un SIG	11
3.2. Analyse statistique	12
3.2.1. Description des variables	12
3.2.2. Analyses statistiques réalisées	12
Partie 3 - Résultats	14
1. Résultats des analyses statistiques univariées	14
2. Modélisation de la survenue de foyers ovins	14
3. Essais de cartes d'aléa de survenue de fièvre catarrhale ovine	15
Partie 4 - Discussion et perspectives	16
1. Discussion sur les données épidémiologiques disponibles et perspectives d'amélioration	16
1.1. Choix des données épidémiologiques	16
1.2. Qualité des données épidémiologiques	17
2. Discussion sur les paramètres environnementaux et nouvelles perspectives de travail	18
2.1. Choix des données environnementales	18
2.2. Choix des méthodes de traitement d'image	18
2.3. Perspectives	19
3. Discussion sur la méthode d'analyse et compléments envisagés	19
4. Discussion des résultats de l'analyse statistique	20
CONCLUSION	22
BIBLIOGRAPHIE	23
GLOSSAIRE	26
ANNEXES	29

LISTE DES FIGURES

<i>Figure 1 : Exemple de schéma permettant de positionner les relevés de végétation</i>	<i>7</i>
<i>Figure 2 : Paramètres de segmentation de l'image SPOT du 06/06/2001</i>	<i>10</i>
<i>Figure 3 : Comparaison des noyaux d'apprentissage selon les deux approches.....</i>	<i>10</i>

LISTE DES TABLEAUX

<i>Tableau 1 : Caractéristiques de l'image SPOT</i>	<i>6</i>
<i>Tableau 2 : Comparaison des tailles des noyaux d'apprentissage et de validation selon les deux méthodes.....</i>	<i>10</i>
<i>Tableau 3 : Résultats des classifications finales selon les deux approches.....</i>	<i>11</i>
<i>Tableau 4 : Variables étudiées pour expliquer l'état sanitaire des élevages.....</i>	<i>12</i>
<i>Tableau 5 : Paramètres du modèle logistique de survenue des foyers ovins ayant comme variables explicatives la mixité, la latitude, les cultures et le NDVI moyen.....</i>	<i>14</i>
<i>Tableau 6 : Paramètres du modèle logistique de survenue des foyers ovins ayant comme variables explicatives la mixité, la latitude, et les classes cultures et maquis-garrigue-forêt</i>	<i>15</i>

LISTE DES ANNEXES

<i>Annexe 1 : Modèles de répartition de C. imicola</i>	<i>30</i>
<i>Annexe 2 : Chronogramme</i>	<i>31</i>
<i>Annexe 3 : Localisation des élevages de Corse du Sud par rapport à l'image satellitale SPOT</i>	<i>32</i>
<i>Annexe 4 : Arbre de décision utilisable pour déterminer le statut sanitaire des élevages ovins, bovins et mixtes</i>	<i>33</i>
<i>Annexe 5 : Fiches de relevés de végétation</i>	<i>34</i>
<i>Annexe 6 : Rapport de mission - Relevés de végétation effectués en Corse du Sud du 13 au 21 août 2004.....</i>	<i>35</i>
<i>Annexe 7 : Attributs choisis et retenus pour la classification orientée objet par plus proches voisins.....</i>	<i>38</i>
<i>Annexe 8 : Procédures d'amélioration des classifications</i>	<i>39</i>
<i>Annexe 9 : Classifications finales.....</i>	<i>40</i>
<i>Annexe 10 : Matrices de confusion des classifications pixel à pixel et orientée objet.....</i>	<i>46</i>
<i>Annexe 11 : Statut sanitaire des élevages ovins de Corse du Sud</i>	<i>47</i>
<i>Annexe 12 : Analyse descriptive et corrélations entre les variables</i>	<i>48</i>
<i>Annexe 13 : Essais de cartes d'aléa de fièvre catarrhale ovine.....</i>	<i>51</i>
<i>Annexe 14 : Poster de vulgarisation scientifique réalisé pour la fête de la science 2004.....</i>	<i>52</i>

Remarques :

Les abréviations explicitées sont notées dans le texte par un dièse (#). Les mots expliqués dans le glossaire sont notés d'un astérisque (*).

LISTE DES ABREVIATIONS

ACP : Analyse en composantes principales

AVHRR : « Advanced very high resolution radiometer »

Cirad : Centre de coopération internationale en recherche agronomique pour le développement

Cirad-emvt : Département élevage et médecine vétérinaire du Cirad

CNRS : Centre national de la recherche scientifique

DGAL : Direction générale de l'alimentation, Ministère de l'agriculture

DDSV : Direction départementale des services vétérinaires

DOM : Département d'outre-mer

Géotrop : Equipe du département amélioration des méthodes pour l'innovation scientifique du Cirad, programme Agronomie

HRV2 : « High resolution visible »

IGN : Institut géographique national

ISIS : Incitation à l'utilisation scientifique d'images SPOT

LST : « Land surface temperature », température de surface

MIR : Moyen infrarouge

MNT : Modèle numérique de terrain

NOAA : « National Oceanic and Atmospheric Administration »

NDVI : « Normalized difference vegetation index », indice de végétation à différence normalisée

NTF : Nouvelle triangulation française

PIR : Proche infrarouge

ROI : « Region of interest », noyau (zone) d'apprentissage ou de validation

SERF : Santé et environnement rural en Franche-Comté

SILAT : Systèmes d'informations localisées pour l'aménagement des territoires

SIG : Système d'information géographique

SPOT : Satellite pour l'observation de la Terre

VPD : « Vapour pressure deficit », déficit de pression de vapeur

Introduction

Du fait des changements climatiques et des transformations induites par les activités humaines, de nombreux écosystèmes méditerranéens ont récemment acquis, ou sont en train d'acquérir, la capacité écologique d'héberger de façon pérenne certains pathogènes majeurs pour la santé des hommes et des animaux, responsables de maladies dites « émergentes* ». Ces maladies se révèlent très fréquemment être à transmission vectorielle. En effet, les vecteurs* étant particulièrement sensibles à leur écosystème, tout changement dans les écosystèmes peut modifier leur répartition et donc celle des maladies qu'ils transmettent.

La télédétection, de par sa capacité à caractériser l'environnement, constitue un outil adapté pour étudier les maladies vectorielles, du fait même du lien existant entre les vecteurs et leur écosystème [1]. Elle a ainsi souvent permis de mieux comprendre l'épidémiologie de maladies vectorielles telles que le paludisme [2; 3; 4] ou la trypanosomose [5; 6; 7].

Nous proposons ici d'utiliser la télédétection pour étudier les conditions d'installation d'un vecteur dans un nouvel écosystème en prenant comme exemple celui de *Culicoides imicola*, le moucheron vecteur de la fièvre catarrhale ovine en Corse. *C. imicola* est un insecte tropical, qui semble s'être installé durablement en Corse dans des écosystèmes très différents de ceux dont il est originaire. Depuis 2000, il a occasionné plusieurs épizooties* dans le bassin méditerranéen ayant entraîné la perte de plus de 300 000 moutons. Le nouvel habitat de ce vecteur reste pour l'instant méconnu. La télédétection pourrait aider à identifier l'environnement qui lui est favorable.

La démarche présentée ici consiste à comparer les paramètres environnementaux caractérisant le voisinage des élevages sains d'une part et infectés d'autre part et à élaborer un modèle expliquant leur répartition.

Dans ce rapport nous présentons tout d'abord le contexte et les objectifs de l'étude (partie 1). Puis les données épidémiologiques, les paramètres environnementaux extraits d'une image SPOT[#] du sud de la Corse et l'intégration de ces données dans un SIG[#] sont détaillées (partie 2). Enfin nous présentons (partie 3) et discutons (partie 4) le modèle obtenu.

Ce projet SILAT[#] est commandité par le Cirad-emvt[#] qui est en charge de la gestion de la crise de la fièvre catarrhale ovine auprès de la DGAL[#]. Ce travail interdisciplinaire est issu d'une collaboration entre plusieurs équipes : le Cirad-emvt pour les aspects d'épidémiologie et d'entomologie, l'équipe Géotrop[#] (Cirad) pour les aspects de traitement d'image et de SIG et l'équipe SERF[#] pour les aspects statistiques.

Le projet SILAT a pour objectif de caractériser par télédétection l'environnement favorable à un vecteur récemment installé dans un nouvel écosystème. Le modèle choisi est celui de *Culicoides imicola*, vecteur de la fièvre catarrhale ovine en Corse.

Partie 1 - Contexte et objectifs

Dans cette première partie, nous présentons la situation de la fièvre catarrhale ovine, puis nous synthétisons les principaux modèles disponibles expliquant la répartition du vecteur dans le bassin méditerranéen. A partir de ce constat, nous analysons les besoins et détaillons les objectifs ainsi que les hypothèses sous-jacentes au travail réalisé.

1. Analyse de l'existant

1.1. Etat des connaissances sur la maladie et le vecteur

La fièvre catarrhale ovine est une maladie virale affectant tous les ruminants sauvages et domestiques, transmise par des moucheron hémato-phages* du genre *Culicoides*. Chez les ovins, la maladie se manifeste par une fièvre, une inflammation et ulcération de la muqueuse buccale, des oedèmes, des raideurs musculaires et des boiteries. En zone nouvellement infectée, elle peut être responsable de la mort de 20 % du cheptel. L'infection est généralement asymptomatique chez les bovins et les caprins mais ces espèces pourraient jouer le rôle de réservoirs* [8]. Il existe 24 sérotypes* de virus qui ne présentent pas de protection croisée*; ainsi un vaccin ne protège que pour le sérotype contre lequel il est dirigé (il existe néanmoins des vaccins regroupant plusieurs sérotypes).

Culicoides imicola est le vecteur de la fièvre catarrhale ovine en Europe. En général, il a une faible dispersion autour de son lieu de reproduction : quelques centaines de mètres. Mais il peut être transporté passivement par les vents sur plusieurs centaines de kilomètres, ce qui expliquerait sa diffusion progressive de l'Afrique sub-saharienne jusqu'à l'Europe [9; 10; 11; 12; 13].

La bio-écologie de *C. imicola* est encore mal connue. Les larves préféreraient des sols riches en matière organique et présentant un niveau d'humidité suffisant. Les femelles adultes, crépusculaires, effectuent plusieurs repas de sang au cours de leur vie, ce qui explique qu'elles puissent s'infecter à partir d'un hôte virémique* puis transmettre le virus à plusieurs autres hôtes. Les adultes sont sensibles à la température (qui influe sur la durée de leur cycle de reproduction, leur survie et leur capacité à transmettre le virus), à l'humidité (qui influe sur leur activité, leur développement et leur survie) et au vent (qui diminue leur activité et permet leur diffusion) [14].

Nous faisons l'hypothèse que les paramètres physiques climatiques nécessaires aux *Culicoides* déterminent des types d'environnements pouvant être identifiés à partir d'images satellitales par l'étude de paramètres tels que le type d'occupation du sol — en particulier la végétation —, l'altitude, l'ensoleillement, l'exposition...

1.2. Modèles existants de répartition du vecteur

Les premiers modèles expliquant la répartition et l'abondance des *Culicoides* s'appuient uniquement sur des variables météorologiques [15; 16; 17]. Depuis peu, les modèles proposés intègrent aux données météorologiques des variables dérivées d'images satellitales telles que le NDVI# (cf. définition dans la partie 2, § 2.2). Les principaux modèles présentés dans la littérature [18; 19; 20; 21; 22] sont synthétisés dans l'annexe 1. Le NDVI est une variable explicative dans les quatre modèles présentés. Il semble être déterminant pour expliquer la présence de *C. imicola*, alors que, pour d'autres espèces de *Culicoides* telles que *C. obsoletus* et *C. newsteadi*, les variables liées à la température seraient les plus importantes [23].

2. Analyse des besoins

Pour l'instant, ces modèles ont tous été développés à partir de données à basse résolution spatiale (NOAA[#] – AVHRR[#]). De plus, aucun n'a été développé en Corse, seul département français (hors DOM[#]) aujourd'hui touché par la fièvre catarrhale ovine. La généralisation dans d'autres pays des modèles développés à l'échelle nationale n'a pas toujours été testée. Certains modèles ont été utilisés pour prédire la situation dans le bassin méditerranéen avec des résultats contradictoires : certains prédisent la présence du vecteur en Corse [22], d'autres non [21]. De plus, ces modèles fonctionnent tous à une échelle relativement petite, de l'ordre du kilomètre. Ils ne peuvent donc expliquer les différences observées au sein même de l'île.

Enfin, bien que la végétation joue vraisemblablement un rôle important dans la répartition des *Culicoides*, rares sont les modèles qui intègrent la notion de paysage [7]. L'approche proposée ici pourrait permettre, grâce aux connaissances acquises via les données satellitales, de mieux appréhender les paramètres écologiques de l'insecte.

La Corse ne dispose pas d'un réseau de pièges assez dense pour nous permettre de proposer un modèle basé sur des abondances de *C. imicola*. Par contre, elle a subi plusieurs épizooties en 2000, 2001, 2003 et 2004. Les données de répartition des foyers* peuvent servir à élaborer le modèle. Nous faisons l'hypothèse que le statut infecté ou indemne des élevages est essentiellement conditionné par un environnement plus ou moins favorable au développement du vecteur. En d'autres termes, les foyers seraient situés dans des zones favorables au moucheron et les élevages sains dans des zones moins favorables.

Un modèle développé à une échelle plus fine et tenant compte de la végétation peut contribuer à expliquer la répartition des foyers en Corse.

3. Objectifs

Les objectifs de ce projet SILAT sont donc de caractériser l'environnement favorable aux *Culicoides* en Corse. Pour cela, un modèle statistique est développé à partir d'une image SPOT afin d'expliquer la répartition des foyers de fièvre catarrhale ovine. Ce modèle permet ensuite d'établir des cartes d'aléa.

Partie 2 - Approche opérationnelle

Dans cette deuxième partie, nous présentons les données utilisées et les méthodes mises en œuvre pour construire le modèle de répartition des foyers de fièvre catarrhale ovine en Corse. Trois phases peuvent être distinguées : l'organisation des données épidémiologiques, puis l'extraction des paramètres environnementaux et enfin l'analyse de l'ensemble de ces données, analyse qui comprend leur intégration dans un SIG et les traitements statistiques.

1. Organisation des données épidémiologiques

Pour créer un modèle de répartition des foyers, les informations suivantes étaient nécessaires : la liste des élevages et leur localisation, leurs caractéristiques (espèces élevées et effectif des troupeaux) et les données sanitaires (élevage infecté/sain) les concernant. Cette phase de recueil, tri et organisation des données est a nécessité beaucoup de temps (cf. annexe 2) du fait de l'hétérogénéité des données disponibles.

1.1. Géoréférencement des élevages ovins et bovins

→ *Méthode*

En France, tout élevage est soumis à déclaration auprès des services vétérinaires départementaux, et chaque animal doit être recensé et identifié. Les services vétérinaires disposent donc de la liste des exploitations par commune. Pour ce projet SILAT, il semblait nécessaire de travailler à une échelle plus grande que l'échelle communale car, au sein d'une commune, certains élevages sont infectés alors que d'autres sont sains. Cette information à un niveau infra-communal avait déjà été recherchée en 2002 [24]. Les élevages ovins et bovins de Corse du Sud avaient été pointés sur des cartes IGN[#] au 1:25 000 par le chef des techniciens vétérinaires de Corse du Sud, M. Bousquet. Nous avons repris ces cartes papier et saisi les exploitations sous ArcInfo[®] en les positionnant à l'aide du scan 25^{®*}. Pour quelques exploitations, les coordonnées étaient déjà connues ; ces dernières ont pu être rajoutées à celles saisies après élimination des doublons. En tout, 404 exploitations ont pu être géoréférencées (cf annexe 3).

→ *Difficultés rencontrées et choix effectués*

L'identifiant unique de ces exploitations est le numéro d'élevage. Mais un même éleveur peut avoir plusieurs exploitations et une exploitation peut appartenir à un ensemble d'éleveurs et/ou concerner des espèces différentes (élevages dits « mixtes »). Enfin, certains troupeaux issus d'exploitations différentes, parfois même appartenant à des propriétaires différents, sont gardés ensemble sur les mêmes pâtures. Malgré cela, nous avons choisi, pour deux raisons, de conserver comme unité l'exploitation en nous basant sur le numéro d'élevage. Tout d'abord, il paraissait difficile et arbitraire de définir une distance minimale pour décider de « fusionner » des élevages. D'autre part, fusionner certaines exploitations aurait pu rendre ces informations inexploitable pour d'autres études à venir. Les raisons invoquées ci-dessus expliquent que certaines exploitations soient localisées au même endroit. La précision de la localisation des exploitations sera discutée dans la partie 4.

1.2. Recueil des caractéristiques sur les élevages

Le recueil des caractéristiques des élevages avait pour but de renseigner chaque exploitation sur les espèces élevées et les effectifs des troupeaux. Aucune base disponible ne réunissait l'ensemble de ces informations. Les données disponibles étaient parcellaires, provenaient d'origines diverses, avaient été recueillies par des personnes différentes sans souci de construire une base unique et ne possédaient pas de méta-données*. Un long travail

de regroupement, de recoupement et de vérification des informations a donc été effectué. Il a nécessité de nombreux échanges (4 réunions ainsi que de nombreux échanges par courriels et par téléphone) avec l'équipe du Cirad-emvt. L'information sur les espèces élevées a pu être complétée pour la majorité des élevages géo-référencés. En revanche, l'information sur la taille du cheptel n'a pu être recueillie de façon assez exhaustive pour être utilisée.

1.3. Recueil de l'information sur l'état sanitaire des élevages

→ *Elevages ovins*

La déclaration de toute suspicion de fièvre catarrhale ovine est obligatoire et permet aux éleveurs ayant des troupeaux atteints de percevoir des aides de l'Etat. Elle engendre une procédure visant à confirmer ou infirmer la suspicion au moyen de tests virologiques. Au terme de cette procédure, les élevages ayant été confirmés positifs par virologie sont officiellement déclarés infectés par l'Agence française de sécurité des aliments. La liste des élevages ovins officiellement infectés a ainsi pu être facilement récupérée pour la période 2000-2003. Tout élevage ovin ne faisant pas partie de la liste a été considéré comme sain.

→ *Elevages bovins*

Les bovins ne manifestent pas de symptômes cliniques, mais peuvent néanmoins être infectés par le virus et développer alors des anticorps décelables par des tests sérologiques. Certains élevages bovins ont fait l'objet de tests sérologiques afin de suivre l'évolution de la maladie, en particulier en 2001-2002, lorsque le nombre de foyers ne permettait plus de traduire le niveau de circulation virale du fait de la vaccination de nombreux élevages ovins.

Initialement, il était prévu d'utiliser les résultats de ces tests sérologiques pour connaître le statut sanitaire d'élevages bovins ou d'élevages mixtes (ovin/bovin), dont les ovins n'auraient pas présenté de symptômes (cf. annexe 4 arbre de détermination du statut prévu initialement). Le nombre d'élevages utilisés et la puissance du modèle construit auraient ainsi pu être augmentés. Mais devant les biais d'échantillonnage des élevages bovins (basé sur le volontariat des éleveurs), ces résultats n'ont pas été exploités. Néanmoins, le caractère mixte ou non des élevages ovins a été conservé, pour tester l'hypothèse de l'existence d'un facteur de risque lié à la présence de bovins (réservoir supposé de la maladie).

Les données concernant l'espèce (§ 1.2) et l'état sanitaire ont été ajoutés à la couche d'élevages géoréférencés en réalisant une jointure sur le numéro d'élevage sous ArcInfo®.

2. Extraction des paramètres environnementaux

2.1. Choix des paramètres environnementaux

Dans cette partie, nous justifions les choix des paramètres environnementaux retenus et présentons les données nécessaires à leur obtention.

→ *Choix effectués*

Le choix des paramètres environnementaux s'est fondé sur les connaissances de la bio-écologie de l'insecte et les modèles présentés dans la littérature. Afin de caractériser l'environnement favorable aux *Culicoides*, il faudrait disposer des données sur la température, l'humidité, le vent et la végétation. Mais il est impossible d'acquérir ces données à une échelle fine. C'est pourquoi nous proposons d'utiliser des indicateurs de ces variables issus de données de télédétection. Les indicateurs retenus sont les suivants :

- **la végétation** : elle résulte d'un ensemble de conditions du milieu : conditions météorologiques, type de sol... Elle peut être caractérisée de façon fine sur une image satellitale au travers des classes de végétation (en réalisant une classification) et au travers d'indices de végétation tels que le NDVI.
- **l'altitude** : elle influe sur les conditions de température.

- **la pente** : elle pourrait conditionner en partie l'humidité du milieu.
- **l'ensoleillement** : il influe sur la température et l'humidité.
- **l'orientation** : elle traduit l'exposition aux vents

Le premier indicateur peut s'obtenir à partir d'images haute résolution, les quatre derniers peuvent être calculés à partir d'un MNT.

→ *Image SPOT*

Tableau 1 : Caractéristiques de l'image SPOT

Une image SPOT du sud de la Corse (cf. annexe 3), datée du 6 juin 2001 avait été acquise grâce au programme ISIS[#] afin de réaliser un premier travail sur la fièvre catarrhale ovine [24]. Cette image, acquise par le satellite SPOT 4, a une résolution spatiale de 10 mètres. Elle résulte de la superposition d'une image acquise en mode monospectral (« M », à 10 mètres) avec une autre image acquise en mode multispectral (« Xi », à 20 mètres). C'est un produit SpotView de niveau 3, orthorectifié (corrigé géométriquement par rapport aux conditions de la prise de vue et projeté en Lambert 4 Carto, méridien de Paris). Les autres caractéristiques de l'image sont présentées dans le tableau 1.

Numéro de l'image	405926601060610161721 Sat+2
Satellite (capteur)	Spot4 (HRV2 [#])
Produit	SpotView (niveau 3 ortho)
Date de prise de vue	06/06/2001
Bandes (µm)	0,50 - 0,59 (vert), 0,61 - 0,68 (rouge), 0,78 - 0,89 (PIR [#]) et 1,58 - 1,75 (MIR)
Taille du pixel	10 mètres (mode multispectral « M+Xi »)
Taille de l'image (Pixels)	7187 lignes x 6566 colonnes (47 189 842 pixels)
Projection (datum, ellipsoïde de référence)	Lambert 4 carto Paris (NTF [#] , ellipsoïde de Clarke 1880 IGN)

Donnée SPOT/ Programme ISIS, © CNES (2001),
distribution Spot Image S.A.

Une deuxième image située au nord de celle-ci a fait l'objet d'une demande auprès du programme ISIS (en date du 12 juillet 2004) afin d'inclure un plus grand nombre d'élevages dans l'étude. Cette demande n'a pas encore aboutie.

→ *MNT*

Le MNT du sud de la Corse a été acquis auprès de l'IGN (BD ALTI[®]). La résolution spatiale (taille du pixel) est de 50 mètres et la précision altimétrique annoncée est décimétrique. Il est projeté en Lambert 4. L'altitude varie de 0 à 1457 mètres sur cette zone.

2.2. Calcul d'un indice de végétation

Le NDVI est un indice de végétation à différence normalisée qui est étroitement corrélé avec l'activité chlorophyllienne des surfaces végétales. Il est basé sur la différence des réflectances mesurées dans le rouge et le proche infrarouge (*Equation 1*) :

$$NDVI = (PIR - Rouge) / (PIR + Rouge).$$

C'est un indicateur qui, lorsqu'il est calculé à basse résolution spatiale, est corrélé à l'abondance de *Culicoides*. Il paraît donc intéressant de l'inclure dans le modèle proposé ici. Il se calcule sous ENVI grâce à la fonction *Transform, NDVI*.

2.3. Réalisation d'une carte de végétation à partir de l'image SPOT

→ *Choix de réaliser des classifications supervisées*

Les classifications « supervisées », contrairement aux classifications « non supervisées », supposent une connaissance préalable du terrain. Le principe est de rechercher les zones qui sont semblables à des zones de référence. Elles impliquent donc de disposer de relevés de végétation afin de constituer des noyaux (ou zones) d'apprentissage et de validation. Les noyaux d'apprentissage sont les classes de référence à partir desquelles le reste

de l'image est classé. Les noyaux de validation servent à estimer le pourcentage de pixels correctement classés. Les classifications supervisées donnent en général de meilleurs résultats.

→ Démarche

En l'absence de relevés de végétation adaptés pour réaliser une classification supervisée, une mission de 7 jours en Corse a été effectuée. Deux types de classification ont ensuite été mis en œuvre. Un premier type de classification a été réalisé selon une méthode classique, dite « pixel à pixel ». Cette méthode ayant donné des résultats moyens, une deuxième approche a été testée. Développée plus récemment, elle s'appuie sur une méthode dite « orientée objet ». Les deux méthodes sont développées ci-dessous (§ 2.3.2. et § 2.3.3.) après présentation de la phase d'acquisition des relevés de végétation (§2.3.1.). Une classification non supervisée (par la méthode « isodata* ») a également été réalisée, afin de comparer ses résultats avec les classifications supervisées. Mais, même avec 50 classes initiales, les confusions étaient très importantes : de nombreuses classes se retrouvaient dans plusieurs noyaux différents. Cette classification n'est donc pas présentée.

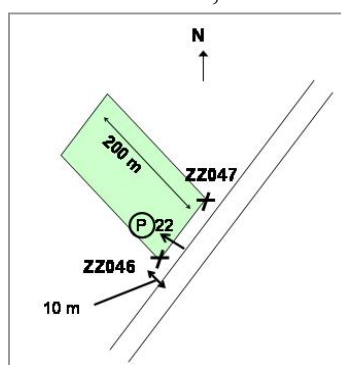
2.3.1. Acquisition des relevés de végétation

→ Objectifs et déroulement de la mission

La mission avait pour objectifs d'effectuer des relevés de végétation dans toute la zone couverte par l'image SPOT ainsi que de prendre contact avec les services vétérinaires départementaux de Corse du Sud et de visiter quelques élevages infectés. Elle s'est déroulée du 13 au 21 août 2004, donc bien après la date d'acquisition de l'image, mais néanmoins à peu près à la même saison. Afin d'effectuer des relevés de végétation sur l'ensemble du pourtour du sud de la Corse, chaque jour, une nouvelle zone géographique était explorée, les déplacements étant effectués en voiture. Pour chaque relevé de terrain, les coordonnées du site étaient relevées par GPS, une fiche de relevé de végétation était renseignée, et des photos de la végétation étaient prises.

→ Mise au point d'une fiche de relevé de végétation

Les classes de végétation auxquelles on voulait aboutir ont été définies en s'appuyant sur les groupes végétaux classiquement décrits en Corse [25] et sur la carte de végétation potentielle du CNRS# (Carte de la végétation de la France : Carte de Corse). Une fois les classes définies, une fiche de relevé de végétation a été élaborée (cf. annexe 5). Un nombre



assez important de classes (28) a été choisi, laissant le choix de les fusionner par la suite. En effet, cette démarche est préférable au choix initial d'un petit nombre de classes, pouvant inclure des sous-classes différentes, et qui conduirait alors à rechercher des classes hétérogènes dans l'image. Les relevés étaient repérés par au moins deux points GPS et un schéma permettant de positionner le noyau (figure 1). Sur ce schéma étaient représentés le nord, la route, la distance de la parcelle relevée à la route, la parcelle, sa taille, les points GPS (ici ZZ046 et ZZ047), l'endroit de prise de vue, le numéro (ici P22) et l'orientation de la photo.

Figure 1 : Exemple de schéma permettant de positionner les relevés de végétation

Les flores [26; 27], qui avaient permis d'apprendre à reconnaître les principales espèces caractéristiques avant le départ en mission, ont été utilisées pour vérifier sur le terrain les cas douteux.

→ Choix des zones de relevés de végétation

Les relevés de végétation doivent être effectués sur des zones répondant au compromis suivant : elles doivent être à la fois les plus grandes et les plus homogènes possibles [28]. Les

zones où effectuer des relevés de végétation ont été pour partie présélectionnées avant la mission et pour partie définies sur le terrain.

Avant la mission, plusieurs documents ont servi à présélectionner des zones homogènes, bien réparties, appartenant à toutes les classes que nous souhaitions retrouver dans la classification :

- Les cartes IGN 1:25000 ont servi à repérer des classes peu fréquentes telles que les marais, à vérifier que les relevés étaient accessibles par la route et à repérer les itinéraires à suivre en voiture ;
- Une composition colorée des trois canaux (contenant le plus de variance) de l'ACP^{##} de l'image SPOT a été produite. L'image ainsi obtenue est plus contrastée que l'image initiale, et il est plus facile d'apprécier les zones homogènes d'un point de vue radiométrique. Des cartes papier ont été établies de façon à pouvoir, à tout moment, se repérer sur le terrain par rapport aux zones ressortant sur l'ACP. Les zones homogènes sur l'ACP et accessibles par la route ont été notées sur les cartes papier afin d'y effectuer éventuellement des relevés de végétation (si ces zones s'avéraient effectivement homogènes sur le terrain).

Sur le terrain, le long des routes, des relevés ont été effectués lorsque des zones homogènes et suffisamment grandes étaient repérées.

→ Bilan de la mission

Cette mission a permis, au final, de réaliser 102 relevés de végétation dans les différentes classes préalablement identifiées (cf. rapport détaillé de la mission en annexe 6).

2.3.2. Classifications

Avant toute exploitation, l'image a tout d'abord été séparée en 2 classes par la méthode « isodata » afin de créer un masque contenant la mer. Cette méthode a donné de meilleurs résultats qu'un masque réalisé à partir des données d'altitude car il y a parfois de larges zones côtières où l'altitude est nulle. Les nuages et leurs ombres ont été découpés manuellement et masqués. La partie masquée de l'image représente 59,7 % de l'image initiale. Cette image « sans mer ni nuage » a fait l'objet des classifications décrites ci-après.

→ *Définition des zones d'apprentissage et de validation*

Pour chaque classe, les relevés de végétation ont été séparés, au hasard, en deux groupes équitables constituant les noyaux d'apprentissage et de validation. En cas de nombre impair de relevés au sein d'une classe, les noyaux d'apprentissage ont été favorisés. A l'aide de l'image et des cartes IGN 1:25000, d'autres noyaux d'apprentissage et de validation ont été saisis pour les classes faciles à repérer : « urbain », « mer », « eau », « sable » et « route ».

→ *Classifications pixel à pixel*

L'approche pixel à pixel est celle qui est la plus couramment utilisée. Elle consiste à effectuer des calculs pour chaque pixel afin de les classer individuellement.

▪ Classification par maximum de vraisemblance

Les classifications par maximum de vraisemblance se basent sur le calcul probabiliste pour chaque pixel d'être rattaché à une classe plutôt qu'à une autre [28]. L'utilisateur est amené à définir un seuil de rejet : il s'agit de la probabilité en deçà de laquelle un pixel p ne sera pas classé dans la classe C_j . Un seuil de 70 % a été choisi afin de diminuer le nombre de pixels non classés et de tenir compte du grand nombre de classes utilisées initialement (certaines devant être fusionnées par la suite).

▪ Evaluation de la qualité des classifications par maximum de vraisemblance

La qualité d'une classification par maximum de vraisemblance s'estime grâce à plusieurs paramètres : les comportements spectraux des noyaux, leur séparabilité*, l'image des

probabilités et la matrice de confusion* [28]. La séparabilité des noyaux peut s'estimer visuellement, préalablement à la classification, grâce à un histogramme pluridimensionnel représentant, pour chaque pixel, la probabilité d'appartenance à tel ou tel noyau. Sous ENVI[®], il est aussi possible de calculer, en prenant les noyaux deux à deux, le coefficient de Jeffries-Matusita, reflétant la séparabilité des noyaux entre eux. Si ce coefficient est inférieur à 1,4, on peut considérer que les noyaux se séparent mal ; s'il est supérieur à 1,7, les noyaux sont bien séparés. L'image des probabilités, obtenue après classification, peut être réalisée pour chaque noyau ou pour tous les noyaux simultanément. Sur ces images, la couleur des pixels est fonction de leur probabilité d'appartenir à la classe considérée. Enfin, la matrice de confusion permet d'estimer le pourcentage de pixels de l'image correctement classés en se référant à une « vérité-terrain » qui peut être une carte déjà existante ou des noyaux de validation. Dans ce dernier cas, le pourcentage des pixels constituant le noyau de validation X qui sont effectivement classés dans la classe X sera calculé.

- Amélioration des classifications par maximum de vraisemblance

Pour améliorer les classifications, certains noyaux ont été fusionnés et certaines classes regroupées lorsque deux types de végétation se confondaient et que leur fusion était cohérente d'un point de vue sémantique. Chaque étape a été testée afin de vérifier si elle avait bien entraîné une amélioration.

Dans quelques cas particuliers, les zones d'apprentissage ou les différentes classes de végétation ont dû être redéfinies. Par exemple, pour les classes « maquis* », « garrigue* » et « pelouse* », plusieurs regroupements ont été testés car, sur le terrain, le découpage présenté sur les cartes du CNRS (ayant servi à définir les classes de végétation) n'a pas été retrouvé. Par ailleurs, pour les prairies notamment, certaines parcelles paraissaient très actives sur l'image (avec une forte valeur dans le PIR) et d'autres beaucoup moins. Cela correspond vraisemblablement au fait que les prairies peuvent être en végétation, voire arrosées, à certaines périodes de l'année et sèches à d'autres. Les relevés étant effectués avec un décalage de 3 ans et 2 mois par rapport à la date de prise de l'image, il est normal de constater de telles différences. Ainsi, à partir de l'image, deux noyaux ont dû être distingués : « prairie sèche » et « prairie en herbe ». Enfin, les noyaux de pelouse ont été redélimités lorsque les relevés de végétation signalaient la présence de roches et lorsque ces noyaux se confondaient effectivement avec la classe roche.

→ *Classifications orientées objet*

- Définition des objets : segmentation

Les classifications réalisées sous le logiciel eCognition[®] sont orientées objet : ce sont des objets et non des pixels qui sont classés. L'image est tout d'abord segmentée en objets homogènes. Un pixel est agrégé au pixel voisin pour lequel l'augmentation d'hétérogénéité est la plus faible, et ce jusqu'à ce qu'un seuil d'hétérogénéité de l'objet, fixé par l'utilisateur, soit dépassé. La segmentation peut être paramétrée grâce à plusieurs facteurs (figure 2). Tout d'abord, l'utilisateur peut choisir un critère d'échelle qui correspond au niveau de finesse de la segmentation, en fonction des objets qu'il souhaite caractériser. Un grand facteur d'échelle se traduira par des objets de grande taille. Pour l'image de Corse du Sud, un facteur d'échelle de 50 a été choisi afin de distinguer des objets tels que les champs et les prairies. Ensuite, l'utilisateur peut définir un critère d'homogénéité qui prend en compte deux facteurs : un facteur spectral et un facteur de forme. Le facteur spectral (radiométrique) a été fixé à 90 %, et le facteur de forme à 10 % car les objets à caractériser se distinguent plus par leur radiométrie que par leur forme (le facteur de forme est aussi paramétrable par un facteur de lissage et un facteur de compacité (*compactness*), respectivement fixés à 80 et 20 %). Le facteur d'échelle semble être celui qui est le plus déterminant [29]. Le choix des autres critères reste relativement empirique.

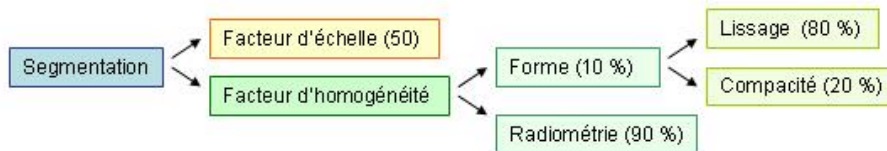
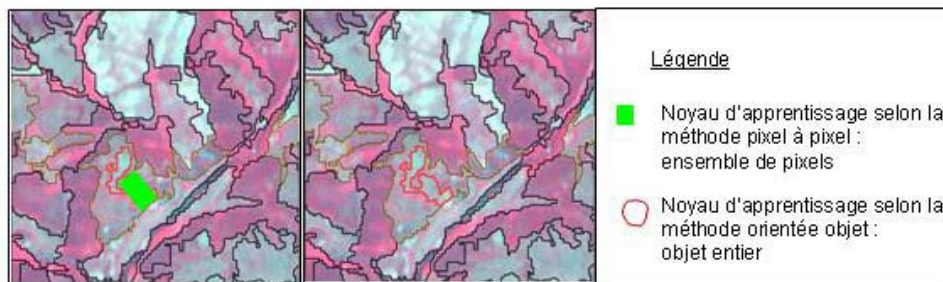


Figure 2 : Paramètres de segmentation de l'image SPOT du 06/06/2001

La classification se base sur des objets d'apprentissage et de validation tout comme ce qui est effectué sous les logiciels proposant une méthode pixel à pixel. Néanmoins, la délimitation des noyaux d'apprentissage et de validation selon les deux méthodes n'a pas les mêmes résultats puisque, avec la méthode orientée objet, les noyaux sont constitués d'objets (issus de la segmentation) qui n'ont pas la même forme que les noyaux délimités avec la méthode pixel à pixel (figure 3).



Donnée SPOT/ Programme ISIS, © CNES (2001), distribution Spot Image S.A.

Figure 3 : Comparaison des noyaux d'apprentissage selon les deux approches

Ainsi, les aires des surfaces d'apprentissage et de validation diffèrent selon les deux méthodes (tableau 2).

Tableau 2 : Comparaison des tailles des noyaux d'apprentissage et de validation selon les deux méthodes

Type de classification	Taille des noyaux d'apprentissage pixels (% de l'image)	Taille des noyaux de validation pixels (% de l'image)
Méthode pixel à pixel	89 925 (0,5)	69 050 (0,4)
Méthode orientée objet	240 778 (1,3)	260 603 (1,4)

▪ Classification par plus proche voisin

Une classification par plus proche voisin a été réalisée. L'utilisateur définit des attributs sur lesquels il souhaite effectuer la classification et un nombre X d'attributs qu'il souhaite conserver pour effectuer la classification (fonction *Feature space optimisation*). Le logiciel crée un espace comportant autant de dimensions qu'il y a d'attributs. Les X attributs permettant d'obtenir la meilleure classification (celle qui permet de séparer au mieux les classes) sont retenus. Chaque objet est alors rattaché à la classe la plus proche de lui dans l'espace multidimensionnel (méthode des plus proches voisins). Les attributs sélectionnés au départ et ceux finalement retenus pour la meilleure classification sont présentés en annexe 7.

▪ Evaluation de la classification par plus proche voisin

Tout comme la classification par maximum de vraisemblance, la qualité de la classification par plus proche voisin peut être estimée en étudiant la séparabilité des noyaux d'apprentissage et en réalisant une matrice de confusion.

▪ Redéfinition des paramètres

Compte tenu du grand nombre de possibilités offertes par le logiciel eCognition[®], mais aussi par manque de temps et de formation, tous les paramètres susceptibles d'améliorer la

classification n'ont pu être testés. Seuls les regroupements de classes et l'influence du nombre de dimensions l'ont été.

→ *Résultats des deux types de classification*

Une synthèse des procédures réalisées selon les deux méthodes est présentée dans l'annexe 8. Les classifications finales retenues sont présentés en annexe 9. Les matrices de confusion des classifications finales sont présentées en annexe 10. Le pourcentage de pixels non classés et la précision totale des meilleures classifications pixel à pixel et orientée objet montrent que l'approche orientée objet a donné de meilleurs résultats (tableau 3).

Type de classification	Pixels non classés	Précision totale
Pixel à pixel	6,9 %	69,4 %
Orientée objet	0,2 %	77,8 %

Tableau 3 : Résultats des classifications finales selon les deux approches

2.4. Extraction des paramètres environnementaux liés à la topographie

A partir du MNT de Corse du Sud, sous ENVI[®], ont été extraits la pente, l'orientation et l'ensoleillement calculé à 16 heures. La pente et l'orientation sont mesurées en angles, la pente variant de 0 à 90° et l'orientation de 0 à 360 ° (0 au nord, mesure dans le sens des aiguilles d'une montre).

3. Analyse

3.1. Intégration des données dans un SIG

Les deux classifications (selon les deux méthodes), l'altitude, la pente, le NDVI, l'orientation et l'ensoleillement ont été intégrés dans un SIG (ArcInfo[®]). Pour caractériser le voisinage de chaque élevage, une zone tampon de 500 mètres de rayon a été créée autour de chaque exploitation. Le choix de cette distance repose sur plusieurs considérations. Tout d'abord, *C. imicola* a une dispersion moyenne de quelques centaines de mètres en général. Ensuite, les élevages ont été pointés à l'endroit où se trouve la bergerie. Nous avons estimé que la plupart des pâtures se trouvent dans un rayon de quelques centaines de mètres d'après la mission de terrain effectuée et les discussions avec l'équipe du Cirad-emvt. Enfin, nous voulions chercher si l'environnement avait une influence à un niveau local.

Pour les classifications, les aires des classes de végétation de la zone tampon ont été extraites (plus exactement, le pourcentage d'aire de chaque classe). Le résultat de la classification orientée objet étant dans un format vecteur (shapefile), le calcul a été effectué grâce à la fonction *Intersect*. Pour la classification selon la méthode pixel à pixel, le résultat étant dans un format raster, la fonction *Tabulate areas* sous ArcView[®] a été utilisée car ArcInfo[®] ne permettait pas de gérer le recouvrement des zones tampons. La même méthode a été mise en œuvre pour le NDVI afin d'extraire (sous Excel[®]) les valeurs minimale, maximale, moyenne, et l'étendue de NDVI ainsi que pour connaître l'orientation majoritaire.

L'orientation majoritaire a été codée en 3 classes équitables — 10-130°, 130-250°, 250-10° (dans le sens des aiguilles d'une montre) — afin de tester par la suite une orientation potentiellement à risque : 130-250° (cf. figure 4). Nous faisons l'hypothèse que cette orientation, qui reflète l'exposition aux vents, pourrait être à risque car elle correspond à l'exposition aux vents venant de Sardaigne, vents qui pourraient transporter *C. imicola* (G. Gerbier, communication personnelle).

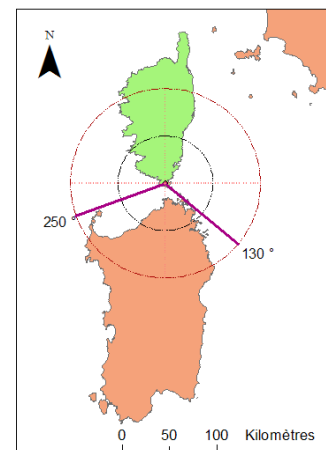


Figure 4 : Orientation des vents venant de Sardaigne vers la Corse

En effet, la première introduction du vecteur pourrait avoir eu lieu par transport passif par les vents en provenance de la Sardaigne. Cela est étayé par le fait qu'à chaque épizootie la Corse a été touchée après la Sardaigne, par les données de piégeages des moucheron, et par le fait que les premiers cas en Corse se déclarent au sud de l'île.

Pour la pente, l'ensoleillement et l'altitude, sous format raster, nous ne pouvions pas calculer la moyenne de chaque zone tampon en utilisant ArcInfo® étant donné leur recouvrement. Un filtre « moyenné » circulaire de 500 m de rayon a donc été créé et appliqué sous ENVI®. Puis la valeur du pixel sur lequel était située l'exploitation a été extraite sous ArcInfo® (fonction *Zonal statistics* de *Spatial analyst*).

3.2. Analyse statistique

3.2.1. Description des variables

- Variables caractérisant les élevages

Le **statut sanitaire** des élevages ovins (n = 89) a été défini en deux classes (indemne ou foyer) (annexe 11). Toutes les épizooties ont été prises entre 2000 et 2003.

Les autres variables caractérisant les élevages sont la **latitude** Y, la **longitude** X et la **mixité** (tableau 4). Cette dernière variable, codée en oui/non, traduit le fait qu'un élevage comporte une (non mixte) ou plusieurs espèces (mixte). Les bovins pouvant jouer le rôle de réservoir de la maladie, il paraissait intéressant de tester si leur présence dans un élevage augmentait le risque d'infection de ce dernier.

- Variables décrivant l'environnement

Les variables décrivant l'environnement peuvent être séparées en deux groupes : celles issues des deux types de classifications et les autres (NDVI, orientation, ensoleillement, pente et altitude) (tableau 4). Un modèle a été construit pour chaque type de classification.

Variables caractérisant les élevages			Foyers (ovins)	X (longitude, en km)	Y (latitude, en km)	mixité	
Variables environnementales	Variables issues des classifications	<i>Aire (en %) des classes (méthode pixel à pixel)</i>	Bâti-sol nu Marais Prairie	Cultures Maquis-garrigue-pelouse-forêt Prairie arborée-Chênes-lièges	Pinède	Eau	Non Classé
		<i>Aires (en %) des classes (méthode orientée objet)</i>	Bâti-sol nu Marais Prairie	Cultures Maquis-garrigue-forêt Prairie arborée-arbres isolés	Pinède	Eau	Non Classé Pelouse
	Autres variables environnementales		NDVI minimum Pente moyenne	NDVI maximum Altitude moyenne	NDVI moyen Ensoleillement moyen	Amplitude NDVI	

Tableau 4 : Variables étudiées pour expliquer l'état sanitaire des élevages

NB 1 : Les variables notées en gras correspondent aux classes n'étant pas exactement similaires.

NB 2 : Les variables sont calculées pour une zone d'un rayon de 500 mètre autour de chaque élevage.

3.2.2. Analyses statistiques réalisées

→ *Analyse descriptive des variables*

Les variables présentées dans le tableau 4 ont fait l'objet d'une analyse descriptive, puis les corrélations des variables 2 à 2 ont été testées. Un test de comparaison de moyenne entre le groupe sain et le groupe infecté a été mis en œuvre pour toutes les variables quantitatives. Un test de Fisher exact a été réalisé pour tester la répartition des variables qualitatives dans le groupe sain et infecté. Ces analyses de statistique univariée ont été réalisées sous SyStat® 10 (SSI, Californie, USA).

→ *Analyse multivariée*

▪ Choix de la régression logistique

L'objectif de cette analyse est de rechercher des variables potentiellement explicatives du statut sanitaire des élevages. La régression logistique est un modèle qui permet d'exprimer la relation entre une variable qualitative à deux classes et des variables qui peuvent être qualitatives ou quantitatives. Il est bien adapté lorsqu'on cherche à expliquer un état de santé qui s'exprime en deux classes : sain/infecté [30]. Le modèle logistique peut être décrit par l'équation suivante (*Equation 2*) :

$$\ln(P/(1-P)) = \alpha + \sum_{i=1}^n \beta_i X_i \Leftrightarrow P = 1/(1 + e^{-\left(\alpha + \sum_{i=1}^n \beta_i X_i\right)})$$

<p><i>P</i> : probabilité d'être infecté <i>α</i> : constante <i>β</i> : coefficient (constant) <i>X</i> : variable explicative</p>

▪ Construction du modèle logistique

Un modèle de régression logistique pas à pas ascendant a été construit sous Egret[®] 2.0 (CYTEL, Cambridge, Mass. USA). Dans un premier temps, les caractéristiques des élevages ont été introduites dans le modèle, puis les variables environnementales ont été proposées à l'entrée du modèle. Une démarche pas à pas ascendante a été mise en œuvre afin de ne retenir que les variables présentant un $p < 0,05$.

▪ Evaluation du modèle

Pour estimer l'adéquation du modèle avec la réalité, la déviance du modèle a été calculée. La déviance, de façon imagée, permet de comparer la réalité observée au travers des données et la réalité telle que le modèle la prédit. Ce qui est prédit par le modèle est proche des données observées si la différence entre ces deux vraisemblances est non significative.

▪ Comparaison des modèles

Puisque les modèles effectués selon les deux méthodes de classification ne sont pas construits à partir des mêmes variables, ils peuvent être comparés en calculant le critère d'Akaike corrigé (Corrected Akaike information criterion, CAIC). Le modèle ayant le plus petit CAIC est le meilleur.

Partie 3 - Résultats

Dans cette partie, nous présentons brièvement les résultats des analyses statistiques effectuées, puis nous détaillons et comparons les modèles logistiques obtenus. Les résultats seront discutés plus précisément en partie 4, § 4. Enfin, nous présentons un exemple de carte d'aléa pouvant être élaborée.

1. Résultats des analyses statistiques univariées

L'analyse descriptive des variables et le tableau des corrélations sont présentés en annexe 12. Le premier constat est que seules les classes « culture », « eau » et « maquis-forêt-garrigue » sont fortement corrélées entre les deux classifications. La variable « Y » (latitude) apparaît comme étant fortement corrélée à plusieurs autres variables : les classes « culture », « prairie » et « non classé » de la méthode pixel à pixel ainsi que le NDVI moyen et maximal. La variable « altitude », et les classes « non classé » et « culture » de la méthode pixel à pixel présentent également des corrélations importantes avec d'autres classes.

2. Modélisation de la survenue de foyers ovins

→ Modélisation de la survenue des foyers ovins effectuée à partir des classes obtenues selon la méthode pixel à pixel

Une fois l'ajustement sur les variables caractérisant les élevages effectué, plus aucune variable n'était retenue dans le modèle.

→ Modélisation de la survenue des foyers ovins effectuée à partir des classes obtenues selon la méthode orientée objet

▪ Modèle 1

Au terme de la démarche pas à pas ascendante, le modèle obtenu est le suivant (avec P : probabilité de survenue d'un foyer) (Equation 3):

$$\ln(P/(1-P)) = C + a \text{ MIXITE} + b Y + c \text{ CULTURES} + d \text{ NDVI moyen}$$

$$\text{Avec } C = 312,9546 \quad a = 3,3704 \quad b = -0,0761 \quad c = -0,0738 \quad d = -7,3115$$

La **déviante** de ce modèle est de **89,48**, son nombre de **degrés de liberté** est **84** et **p = 0,32**. Les détails relatifs aux paramètres du modèle sont synthétisés dans le tableau 5.

Tableau 5 : Paramètres du modèle logistique de survenue des foyers ovins ayant comme variables explicatives la mixité, la latitude, les cultures et le NDVI moyen

Variables	Coefficient	Ecart type	valeur de p	Odds Ratio	
				Valeur	Intervalle de confiance (95 %)
MIXITE	3,37	1,28	0,008	29,09	2,38 - 356,17
Y	-0,08	0,03	0,020	0,93	0,87 - 0,99
CULTURES	-0,07	0,03	0,007	0,93	0,88 - 0,98
NDVI MOYEN	-7,31	3,70	0,048	0,00	0 - 0,94

▪ Modèle 2

Le NDVI étant très corrélé avec la classe « maquis-garrigue-forêt », un deuxième modèle peut être élaboré en remplaçant le NDVI par cette classe (Equation 4) :

$$\ln(P/(1-P)) = C + a \text{ MIXITE} + b Y + c \text{ CULTURES} + d \text{ MAQUIS/GARRIGUE/FORET}$$

$$\text{Avec } C = 415,5859 \quad a = 3,7087 \quad b = -0,1004 \quad c = -0,0901 \quad d = -0,0408$$

La **déviante** de ce modèle est de **85,45**, son nombre de **degrés de liberté** est **84** et **p = 0,44**. Les détails relatifs aux paramètres du modèle sont synthétisés dans le tableau 6.

Tableau 6 : Paramètres du modèle logistique de survenue des foyers ovins ayant comme variables explicatives la mixité, la latitude, et les classes cultures et maquis-garrigue-forêt

Variables	Coefficient	Ecart type	valeur de p	Odds Ratio	
				Valeur	Intervalle de confiance (95 %)
MIXITE	3,71	1,42	0,009	40,80	2,51 - 662,50
Y	-0,10	0,03	< 0,001	0,90	0,85 - 0,96
CULTURES	-0,09	0,03	0,002	0,91	0,86 - 0,97
MAQUIS GARRIGUE FORET	-0,04	0,02	0,008	0,96	0,93 - 0,99

→ Comparaison des deux modèles : critère d'Akaike corrigé

La valeur du **critère d'Akaike corrigé** pour le premier modèle est de **100,16** et de **96,16** pour le second.

3. Essais de cartes d'aléa de survenue de fièvre catarrhale ovine

Des essais de cartes d'aléa réalisés à partir du premier modèle sont présentés en annexe 13 afin d'illustrer les exploitations possibles du modèle. Ces cartes ont été réalisées en inversant les modèles obtenus et en calculant la probabilité de survenue d'un foyer pour chaque pixel. Pour cela, à partir de la classification orientée objet, le pourcentage de classes de cultures et la moyenne du NDVI dans un rayon de 500 mètres ont été calculés pour chaque pixel. Puis, sous ERDAS[®], grâce à la fonction *Modeler, Model maker*, la probabilité par pixel a été calculée à partir de l'équation 2 et 3. Comme il n'est pas possible d'attribuer une valeur de « mixité » aux pixels, deux cartes ont été réalisées : une pour les élevages mixtes et une pour les élevages non mixtes. Ces essais de cartes d'aléa sont à interpréter avec beaucoup de précautions (cf. partie 4, § 4) et ne sont présentés qu'à titre d'illustration.

Partie 4 - Discussion et perspectives

Dans cette partie, nous discutons le travail effectué, en commençant par la qualité des données épidémiologiques disponibles. Ensuite, des propositions sont faites concernant le choix des paramètres environnementaux à inclure dans le modèle. Enfin, les apports et limites du modèle ainsi que les précautions avec lesquelles il convient d'interpréter les cartes d'aléa sont détaillés. Les perspectives apportées par ce travail sont précisées au fur et à mesure pour chacune des étapes.

1. Discussion sur les données épidémiologiques disponibles et perspectives d'amélioration

L'épidémiologie de la fièvre catarrhale ovine est complexe de par l'installation récente du vecteur — peu connu — et des stratégies de lutte mises en œuvre (campagnes de vaccination) qui rendent le suivi de la maladie difficile. Il est ainsi très important de discuter de la qualité des données épidémiologiques car la fiabilité des résultats en dépend. Pour cette étude rétrospective, le choix a souvent été guidé par la disponibilité des données (cf. partie 2 §1), mais des perspectives d'amélioration sont proposées.

1.1. Choix des données épidémiologiques

L'objectif de ce travail est de caractériser l'environnement favorable aux vecteurs de la fièvre catarrhale ovine en Corse. De nombreuses possibilités étaient envisageables : travailler sur les vecteurs ou la maladie, sur les ovins ou les bovins, et enfin sur quelles épizooties.

→ *Vecteur ou maladie ?*

Les données sur les vecteurs restent difficiles à obtenir pour plusieurs raisons : le piégeage requiert des moyens humains importants sur le terrain, et l'identification des insectes qui s'en suit nécessite un long travail de spécialiste (parfois jusqu'à une semaine pour un seul piège). Les données disponibles sur les vecteurs proviennent de 12 points de piégeages sur l'ensemble de la Corse, relevés toutes les trois semaines depuis février 2002. Elles pourraient faire l'objet d'un travail sur les variations temporelles d'abondance en fonction des conditions météorologiques, mais elles ne sont pas en nombre suffisant pour envisager une approche spatiale sur l'identification d'environnements à risque. Le choix s'est donc reporté sur les données concernant la maladie et non le vecteur. La maladie peut être considérée comme un indicateur de la présence du vecteur. Mais l'absence de maladie ne signifie pas l'absence de vecteur. En effet, pour que la maladie apparaisse, il faut que trois conditions soient réunies : la présence du virus, du vecteur et des hôtes (les moutons ici). C'est une première limite de ce travail : il ne permet de détecter que des conditions environnementales favorables à la fois au vecteur et aux moutons. Néanmoins, ce travail permet d'élaborer les premières hypothèses sur l'environnement favorable au vecteur. Favorables au vecteur et à ses hôtes, ces zones sont véritablement à risque et leur connaissance permettrait d'améliorer la surveillance sanitaire.

→ *Ovins ou bovins ?*

Le choix a été fait de travailler uniquement sur les ovins à cause des biais d'échantillonnage des élevages bovins. Néanmoins la présence de bovins a été prise en compte au travers de la variable mixité comme facteur de risque éventuel.

→ *Quelles épizooties prendre en compte ?*

Les données de toutes les épizooties (2000, 2001 et 2003) ont été prises en compte afin de détecter le plus large éventail de zones potentiellement favorables au vecteur et de réduire la probabilité de rencontrer des zones que le moucheron n'a pas eu le temps de coloniser ou bien encore où il est présent mais où le virus ne circule pas.

De plus, les données de chaque épizootie prise individuellement présentent des biais différents. La répartition de la maladie a été conditionnée en 2000 par une propagation encore limitée du vecteur nouvellement introduit par la Sardaigne, en 2001 par la vaccination d'une partie des ovins de Corse et en 2003 par une diffusion limitée d'un nouveau sérotype (sérotype 4), en provenance sans doute de Sardaigne. Le choix a été fait de travailler indistinctement sur l'ensemble de ces données dans l'espoir d'ainsi diluer les biais propres aux données de chaque épizootie.

1.2. Qualité des données épidémiologiques

→ Localisation des élevages

Le géoréférencement des élevages soulève de nombreux problèmes. Tout d'abord, l'inventaire réalisé par M. Bousquet n'est pas exhaustif : sur les quelques 1500 élevages de Corse du Sud, seuls 404 ont pu être géoréférencés. De plus, ce travail a été réalisé en 2001 ; il faudrait tenir compte des changements intervenus depuis (cessations, nouveaux éleveurs...). L'information obtenue auprès de M. Bousquet reste néanmoins très précieuse : elle n'existe que par sa connaissance du terrain et n'est pas répertoriée ailleurs (seule la distribution des élevages à l'échelle de la commune est connue par les DDSV).

Par ailleurs, la localisation des élevages se limite à l'emplacement de la bergerie alors que les animaux sont le plus souvent au pâturage pendant la période à risque. Pour connaître les véritables lieux de transmission, il faudrait donc disposer du parcellaire agricole. Ces données ne sont pas disponibles et leur obtention représenterait une charge de travail très importante. Elle n'est raisonnablement envisageable que pour une zone limitée.

Une autre information cruciale manque : les animaux ont-ils transhumé ? En effet, les animaux transhumants ont sans doute moins de chance d'être infectés car les conditions en altitude sont moins favorables aux *Culicoides* (en effet, ils ne seraient plus retrouvés au delà de 900 mètres d'altitude dans le bassin méditerranéen). L'absence de cette information biaise donc l'étude dans le sens où des zones d'élevage considérées comme indemnes peuvent être favorables aux *Culicoides*, mais, les animaux n'y étant pas présents pendant la saison de transmission, ils ne sont pas à l'origine de foyers. En revanche, cela a sans doute peu d'influence sur la véracité de l'information sur les foyers.

Ces informations sur le parcellaire et les éventuelles transhumances semblent néanmoins essentielles pour définir précisément les zones à risque. Pour les obtenir, une nouvelle enquête de terrain auprès de M. Bousquet, des vétérinaires sanitaires et des éleveurs pourrait être envisagée sur une zone limitée. Il faudrait aussi connaître les dates et les lieux de transhumance pour savoir si les animaux ont quand même pu être exposés ou non. Or ces informations paraissent difficile à obtenir rétrospectivement au vu du temps écoulé ; il faudrait donc se concentrer sur une épizootie récente.

→ Caractéristiques de l'ensemble des élevages ovins et bovins

Les données escomptées au départ, c'est-à-dire la liste et l'effectif de tous les élevages d'ovins et de bovins (voire de caprins) de Corse du sud, n'ont pas pu être recueillies de façon exhaustive. Un travail en collaboration avec la DDSV[#] de Corse du Sud s'impose afin de constituer avec eux un fichier regroupant les informations nécessaires sur les élevages. Il devrait être possible de recueillir ces données pour l'année en cours. En revanche, pour les années précédentes, cela ne pourra être possible que si des sauvegardes annuelles des bases de données ont été effectuées entre 2000 et 2003. Les bases qui pourraient ainsi être obtenues devront être recoupées afin de recenser les exploitations en activité entre 2000 et 2003 (voire 2004). L'absence de données sur l'effectif des élevages a lourdement pénalisé ce travail (cf. paragraphe suivant). Si un tel recueil peut être effectué, il faudra aussi envisager le géoréférencement des « nouveaux » élevages recensés.

→ Hétérogénéité des données sanitaires

Les données sanitaires sur les élevages ovins paraissent relativement fiables, notamment parce que l'indemnisation incite les éleveurs à déclarer la maladie. De plus, la confirmation par virologie des élevages suspects garantit une meilleure fiabilité que les seuls tests sérologiques [31].

En revanche, les données sérologiques des bovins ne sont pas représentatives car elle sont ni exhaustives ni échantillonnées de façon aléatoire mais plutôt basées sur le volontariat des éleveurs. De plus, au sein même des troupeaux prélevés, tous les animaux n'ont pas été testés, et la proportion d'animaux testée est le plus souvent inconnue ou alors très variable. Or, plus on teste d'animaux, plus on a de chance de trouver des séropositifs. Par ailleurs, on ne peut être certain du statut indemne d'un élevage que si tous les animaux ont été testés. C'est pourquoi les modèles logistiques n'ont pas pris en compte les données sur les bovins.

Pour l'analyse statistique, seuls les élevages dont les caractéristiques et le statut était connus ont été inclus. La question de la répartition des données renseignées par rapport à celle des données manquantes mériterait d'être testée, afin de vérifier si les données dont nous disposons ne sont pas agrégées.

Les données épidémiologiques présentent donc potentiellement de nombreux biais mais, en tenant compte des suggestions exposées, elles méritent d'être exploitées.

2. Discussion sur les paramètres environnementaux et nouvelles perspectives de travail

2.1. Choix des données environnementales

→ *Choix des données de télédétection*

Une des nouveautés de ce travail réside dans le choix de travailler à haute résolution spatiale puisque tous les modèles de répartition des *Culicoides* existants ont été établis à partir de données à basse résolution. Ce travail permet d'élaborer des premières hypothèses sur les zones à risque, qui devront être vérifiées et affinées pour aider à orienter les études entomologiques et écologiques réalisées sur le terrain. Il constitue aussi une première approche pour aborder un travail sur la structure du paysage favorable aux *Culicoides*.

→ *Choix des paramètres environnementaux*

Les paramètres environnementaux choisis ont l'avantage d'être d'acquisition facile. Ce sont des indicateurs des trois facteurs bio-écologiques importants pour le vecteur : température, humidité et vent. Ils pourront être complétés par d'autres indicateurs (cf. § 2.3).

2.2. Choix des méthodes de traitement d'image

Ce projet a permis de tester l'apport de deux types de classification très différents. Les résultats des classifications pixel à pixel sont assez moyens ; ils ont pu être légèrement améliorés au terme d'un long travail de regroupement, fusion et redéfinition des classes. Le nombre de relevés de végétation utilisés pour définir les noyaux d'apprentissage est pourtant relativement élevé. Les résultats issus des matrices de confusions sont confortés par l'analyse visuelle des classifications. En effet, la classification pixel à pixel présente de nombreuses confusion, notamment pour les marais (qui se confondent avec le maquis) et les cultures (qui se confondent avec les sols nu ou l'urbain).

La classification orientée objet a permis d'obtenir de meilleurs résultats pour plusieurs raisons. Tout d'abord, les noyaux d'apprentissage utilisés dans cette méthode sont plus grands. Il est possible de découper les objets issus de la segmentation (fonction *Image object cutting*) afin de redessiner exactement la zone relevée sur le terrain. Cela pourrait être testé pour comparer les deux méthodes, mais tel n'était pas l'objectif ici. C'est pourquoi, le plus souvent, nous avons conservé les zones homogènes issues de la segmentation. La finesse de la segmentation mise en œuvre paraît adéquate : le voisinage des élevages comportent plusieurs classes différentes, les parcelles cultivées et de prairie sont reconnues. Une autre raison

expliquant les meilleurs résultats de l'approche orientée objet est le grand nombre et la variété des attributs pouvant servir à réaliser la classification. Ces attributs tiennent compte de la radiométrie mais aussi de facteurs texturaux et de paramètres relatifs à la forme des objets. Les attributs permettant la meilleure discrimination des classes sont retenus. Parmi les attributs retenus ici, deux sont liés à la texture (cf. annexe 7). Ces résultats renforcent l'intérêt de travailler sur la texture. Ensuite, le seuil de classification utilisée dans la méthode orientée objet est plus bas, expliquant le faible pourcentage de zones non classées. Enfin, le type d'approche orientée objet repose sur une méthode propre, qui pourrait s'avérer, dans les années à venir, plus puissante que l'approche classique, pixel à pixel. Compte tenu des résultats en termes de classification et d'apport pour la modélisation, les classifications orientées objet mériteraient d'être explorées de façon plus approfondie.

2.3. Perspectives

→ Acquisition d'une nouvelle image

Une deuxième image a fait l'objet d'une demande (juillet 2004) auprès du programme ISIS afin d'inclure un plus grand nombre d'élevages de Corse du Sud. Si cette demande est acceptée, elle permettra de construire un modèle plus puissant. Une image de Haute Corse pourrait servir à valider le modèle, une fois les élevages de cette région géoréférencés. Afin d'appliquer le modèle sur une autre image, il faudra appliquer des corrections radiométriques (afin de calculer les luminances*, puis, en tenant compte des paramètres d'ensoleillement, les réflectances*) et atmosphériques. Une fois validé, le modèle pourrait être exporté, notamment sur les côtes continentales françaises, afin de prédire les zones où le risque d'apparition de la fièvre catarrhale ovine est le plus fort et les surveiller en priorité. Un travail pourra également porter sur la robustesse du modèle lors du processus de changement d'échelle. Ces travaux sont envisagés dans le cadre d'une thèse poursuivant ce premier travail de maîtrise.

→ Intégration de nouveaux paramètres environnementaux

La présence de petits ruisseaux, relevés sur les cartes IGN 1:25 000 mais non discernables sur l'image satellitale, pourrait constituer un complément très intéressant à exploiter dans le modèle. Le modèle pourrait également être couplé avec des données de télédétection basse résolution telles que la température de surface, la température de l'air, le déficit de pression de vapeur... Cela nécessiterait alors de nouvelles acquisitions d'images.

Une approche paysagère doit aussi être envisagée. A notre connaissance, cette voie n'a pas encore été décrite dans la littérature. Elle permettrait d'obtenir des descriptions plus écologiques de l'habitat des vecteurs. Cette approche pourrait être mise en œuvre selon deux procédures : intégration des paramètres de texture dans les classifications pixel à pixel ; classification fine, sous Clapas[®], des paysages situés dans le voisinage des élevages.

→ Amélioration des techniques de classification

L'approche pixel à pixel pourrait être améliorée en prenant en compte des paramètres de texture (en réalisant par exemple une ACP sur les néo-canaux issus de la matrice de co-occurrence et en combinant les néo-canaux portant le plus d'information aux canaux initiaux de l'image afin de réaliser une nouvelle classification). L'approche orientée objet offre de très nombreuses possibilités (notamment les paramètres utilisés pour caractériser les classes) qui mériteraient d'être explorées.

3. Discussion sur la méthode d'analyse et compléments envisagés

→ Extraction des données à partir d'un SIG

A défaut de connaître le parcellaire des élevages, une zone tampon de 500 mètres a été créée autour des bergeries afin d'étudier l'environnement dans le voisinage des élevages. Le choix de cette distance a été justifié auparavant (Partie 2, § 3.1.). Cette étape a mis en évidence certaines difficultés d'ordre technique (l'extraction des paramètres issus de couches au format raster n'a pas pu être faite aisément en utilisant les fonctions de Spatial Analyst à

cause du recouvrement de certaines zones tampons) et la nécessité de mettre en place une procédure automatique d'extraction des données. Cette manipulation étant amenée à se répéter (à partir de données issues d'une deuxième image de Corse du Sud et/ou d'une image de Haute Corse), il est envisagé d'écrire un script permettant d'extraire toutes les informations provenant des différentes couches pour chaque zone tampon, et d'étendre par itération les extractions à chaque zone tampon.

→ *Analyse statistique*

Le modèle de régression logistique est classiquement employé en épidémiologie. Il n'a été possible de le mettre en œuvre uniquement pour les ovins. A moins de mener une nouvelle enquête sérologique dans les élevages bovins, il paraît difficile de proposer un modèle pour cette espèce. Pourtant, cela pourrait permettre d'identifier de nouvelles zones à risque (si le modèle diffère de celui basé sur les ovins); ces zones pourraient correspondre à un environnement favorable au moucheron et aux bovins mais non aux ovins. Ces zones ne seraient véritablement « à risque » que dans le cas d'introduction d'un cheptel ovin ou si la présence dans les environs de bovins infectés est identifiée comme un facteur de risque. Leur identification peut être considérée comme un objectif de second ordre, moins urgent et moins important que la caractérisation de l'environnement favorable aux *Culicoides* et aux ovins.

Il faudra compléter l'analyse statistique en recherchant une auto-corrélation* spatiale. En effet, le recouvrement de certaines zones tampons implique la corrélation des variables mesurées dans le voisinage des élevages. Il faudrait donc tester leur l'auto-corrélation.

4. Discussion des résultats de l'analyse statistique

Les deux modèles présentés ont une déviance non significative (partie 3, § 2); tous deux représentent donc bien la réalité (telle qu'elle est observée au travers des données).

→ *Les facteurs de risque et les facteurs protecteurs identifiés par les modèles*

Les modèles proposés permettent d'identifier plusieurs variables explicatives significatives. Nous discutons le caractère explicatif de ces variables et les biais potentiels.

La **latitude** paraît être un élément déterminant pour expliquer la répartition des élevages ovins infectés : plus les élevages sont situés au sud de l'île, plus ils sont à risque. Cela pourrait s'expliquer de différentes façons. Tout d'abord, la proximité avec la Sardaigne pourrait être un facteur de risque non négligeable. Ce résultat tend donc à conforter l'hypothèse selon laquelle des moucherons infectés pourraient être transportés par les vents en provenance de Sardaigne. D'autre part, le fait d'avoir inclus les données de 2000, alors que le moucheron n'avait pas encore diffusé jusqu'en Haute Corse, pourrait biaiser le modèle en augmentant l'importance de la latitude, mais l'effet latitude, très significatif ($p = 1 \%$ dans le modèle 1 et $p < 1 \%$ pour le modèle 2), n'est sans doute pas lié uniquement à la prise en compte des données de 2000. Enfin, un autre biais potentiel pourrait être lié à une éventuelle répartition des élevages à fort effectif vers le sud; c'est une autre raison pour laquelle il est indispensable d'obtenir cette information.

La **mixité** des élevages apparaît comme un facteur de risque majeur dans les deux modèles (avec un « Odds ratio* » de 29 pour le modèle 1 et de 40 pour le modèle 2). Cela peut s'expliquer par le fait que la présence de bovins, en jouant le rôle de réservoirs*, pourrait augmenter le risque d'infection des moutons. (En revanche, les résultats d'éventuelles analyses sérologiques sur les bovins ne sont pas pris en compte pour la définition d'un « foyer »; ainsi, ce biais direct a été évité.) Si les élevages mixtes sont aussi les élevages ayant des effectifs ovins les plus importants, le facteur mixité serait alors biaisé puisqu'un élevage plus important a plus de chance d'être infecté. Cela mériterait d'être vérifié car, si l'hypothèse de biais peut être infirmée, cela impliquerait directement les bovins, et les mesures sanitaires telles que les traitements des animaux par des insecticides devraient être étendues aux bovins.

La proportion de la classe « **culture** », qui regroupe les vignes et les céréales, dans un rayon de 500 mètres autour des élevages apparaît comme un facteur protecteur. Plusieurs hypothèses peuvent être formulées pour expliquer ce résultat. Tout d’abord, la présence de cultures pourrait refléter les zones d’élevage intensif, zones qui se caractérisent peut-être par certaines conduites d’élevages (distribution d’aliments, pratiques sanitaires, types de bâtiments...). Une deuxième hypothèse pourrait être que le traitement de ces cultures par des insecticides pourrait avoir une influence sur les vecteurs. Enfin, les zones cultivées, pourraient ne pas être favorables à *C. imicola* en étant exposées au vent et à la chaleur et en ne présentant pas ou peu de gîtes larvaires et/ou d’endroits abrités pouvant servir de gîtes de repos.

Comme d’autres études, le modèle 1 montre que le **NDVI**, qui caractérise la végétation, semble avoir un rôle important sur la distribution de la maladie. Les résultats obtenus ici ne sont toutefois pas directement comparables avec ceux de la littérature car ces derniers, réalisés à partir de données basse résolution, se basent sur les variations annuelles du NDVI et non sur les valeurs du NDVI. Dans ce modèle, le NDVI apparaît comme un facteur protecteur, mais les raisons pouvant expliquer pourquoi le NDVI est protecteur restent à déterminer. Pour cela, il faut explorer plus finement quels sont les paysages protecteurs parmi les zones à fort NDVI.

Le modèle 2 montre qu’il faut rechercher les paysages protecteurs parmi ceux qui présentent du maquis, de la garrigue ou de la forêt de feuillus. Ces résultats concernant la classe « **maquis-garrigue-forêt** » — mais aussi la classe « culture » — constituent un premier pas vers la description écologique des zones favorables et défavorables au moucheron. Elle précise aussi la notion de zone protectrice, auparavant définie comme « zone où le NDVI est élevé ». Il faudrait donc retravailler plus précisément sur ces classes. Cette approche devra être ré-itérée si le modèle, une fois validé, doit être exporté hors de la Corse, puisque le maquis est une formation végétale spécifique à la Corse.

→ *Comparaison des deux modèles*

Les deux modèles sont intéressants, en termes statistiques (déviante faible) et sémantiques. Le modèle basé sur le NDVI a l’avantage d’être plus directement exportable, mais le modèle 2 se rapproche plus d’une description écologique des paysages. Le critère d’Akaike et la valeur déviante sont plus en faveur du second mais les deux restent proches.

Dans la suite de la démarche, il conviendra de valider les modèles à partir d’un nouveau jeu de données avant de les exporter. Cela pourrait être fait à partir d’une deuxième image de Corse du Sud (située plus au nord que la première) ou de Haute Corse.

→ *Essai de carte d’aléa*

Des cartes d’aléa ont été réalisées afin d’illustrer le type de résultat que l’on peut obtenir à partir d’un modèle logistique. Elles doivent être interprétées avec encore plus de précautions que le modèle puisque le modèle permet de rechercher des facteurs expliquant la survenue de foyers et que les cartes d’aléa généralisent les résultats du modèle à l’ensemble d’une zone. C’est pourquoi elles ne sont présentées que très rapidement.

Les biais, présentés au fur et à mesure de la discussion, sont essentiellement liés à la qualité des données initiales disponibles et non au travail réalisé. Bien que les modèles statistiques semblent représenter correctement la réalité telle qu’elle est observée au travers des données, ils doivent être considérés avec précaution. Ils devront être validés avant toute autre exploitation. De nombreuses perspectives sont présentées afin d’améliorer et de compléter ce travail. Elles seront mises en œuvre dans le cadre de la thèse qui suit ce projet de maîtrise.

CONCLUSION

Ce projet a permis de mettre en œuvre la démarche proposée initialement : construire un modèle expliquant la répartition des foyers de fièvre catarrhale ovine en Corse et établir ensuite des cartes d'aléa. Mais les données épidémiologiques n'étaient pas assez complètes pour élaborer un modèle solide. Les nombreux biais potentiels des modèles ont été présentés dans la partie 4 et les cartes d'aléa ne sont donc présentées qu'à titre d'illustration.

Nous espérons que la présentation de la démarche complète auprès des différentes personnes impliquées dans la gestion de la fièvre catarrhale ovine permettra d'ouvrir de nouvelles collaborations, notamment avec la DDSV de Corse du Sud, collaborations indispensables pour construire un modèle à partir de données épidémiologiques homogènes, validées et non biaisées.

Les apports de ce projet en termes de nouvelles perspectives de recherche sont importants. Comme cela a été détaillé dans la partie discussion, de nouvelles pistes sont à explorer à tous les niveaux, que ce soit dans les méthodes de traitements d'image (tester les apports de la texture dans la classification pixel à pixel, approfondir les possibilités offertes par l'approche orientée objet, explorer une approche plus paysagère — effets de lisère, de corridor, de continuité, de contiguïté — notamment sous Clapas[®]), dans l'analyse statistique (re-tester les facteurs de risque afin d'éliminer d'éventuels facteurs de confusion...) ou en matière de validation. En effet, le nouveau modèle ainsi obtenu devra être validé et exporté en territoire indemne afin de répondre à la demande de surveillance des côtes continentales françaises. Ceci est devenu particulièrement urgent puisque de récents piégeages (septembre 2004) réalisés dans le Var ont montré pour la première fois qu'une population de *C. imicola*, s'était introduite sur le continent français.

En tant que projet SILAT, ce travail nous a permis d'aborder un large éventail de domaines de compétences. Il a impliqué des formations dans des champs très variés tels que la botanique et l'approche orientée objet. Il a fait appel aux techniques de traitement d'image (classifications, géoréférencement, filtrage), de SIG (analyse spatiale, cartographie) et de bases de données (organisation des données épidémiologiques). Il nous a également permis de mettre en œuvre les techniques de conduite de projet : de nombreux contacts ont été pris avec les personnes impliquées autour du projet (équipe Cirad, DDSV, éleveurs) ; au niveau organisationnel, une mission de terrain a été préparée et menée à terme, et le travail à venir (qui sera réalisé dans le cadre d'une thèse) a été préparé (demande d'une nouvelle image, propositions d'améliorations à apporter au travail réalisé). Enfin, en termes de communication scientifique, ce projet a fait l'objet d'une synthèse bibliographique, d'un poster de vulgarisation pour la fête de science (présenté en annexe 14), de ce mémoire de master, d'un mémoire de conduite de projet, et de présentations orales auprès des équipes participantes.

Ce travail constitue donc une étape indispensable qui a permis de dégager de premières hypothèses et proposer de nombreuses perspectives qui seront reprises dans le cadre de la poursuite de ce projet par un travail de thèse.

BIBLIOGRAPHIE

- 1 . Tran, A. (2004)
Télédétection et épidémiologie : Mise au point d'indicateurs environnementaux et de modèles dynamiques pour la surveillance des maladies transmissibles. Cas de la dengue en Guyane. Strasbourg, Université Louis Pasteur (Strasbourg I), 205 p.
- 2 . Beck, L. R., M. H. Rodriguez, S. W. Dister, A. D. Rodriguez, E. Rejmankova, A. Ulloa, R. A. Meza, D. R. Roberts, J. F. Paris, M. A. Spanner, R. K. Washino, C. Hacker et L. J. Legters (1994).
Remote Sensing as a landscape epidemiologic tool to identify villages at high risk for malaria transmission. *Am J Trop Med Hyg* **51** (3): 271-280.
- 3 . Hay, S. I., R. W. Snow et D. J. Rogers (1998).
From predicting mosquito habitat to malaria seasons using remotely sensed data : practice, problems and perspectives. *Parasitology Today* **14** (8): 306-312.
- 4 . Omumbo, J. A., S. I. Hay, S. J. Goetz, R. W. Snow et D. J. Rogers (2002).
Updating historical maps of malaria transmission intensity in East Africa using remote sensing. *Photogrammetric Engineering & Remote Sensing* **68** (2): 161-166.
- 5 . Rogers, D. J., G. Hendrickx et J. H. Slingenbergh (1994).
Tsetse flies and their control. *Rev Sci Tech* **13** (4): 1075-124.
- 6 . De La Rocque, S., J. F. Michel, G. De Wispelaere et D. Cuisance (2001).
[New tools for the study of animal trypanosomiasis in the Sudan: model-building of dangerous epidemiological passage by remote sensing geographic information systems]. *Parasite* **8** (3): 171-95.
- 7 . De La Rocque, S., V. Michel, D. Plazanet et R. Pin (2004).
Remote sensing and epidemiology: examples of applications for two vector-borne diseases. *Comp Immunol Microbiol Infect Dis* **27** (5): 331-41.
- 8 . MacLachlan, N. J. (1994).
The pathogenesis and immunology of bluetongue virus infection of ruminants. *Comp Immunol Microbiol Infect Dis* **17** (3-4): 197-206.
- 9 . Sellers, R. F., D. E. Pedgley et M. R. Tucker (1978).
Possible windborne spread of bluetongue to Portugal, June-July 1956. *J Hyg (Lond)* **81** (2): 189-96.
- 10 . Sellers, R. F., E. P. Gibbs, K. A. Herniman, D. E. Pedgley et M. R. Tucker (1979).
Possible origin of the bluetongue epidemic in Cyprus, August 1977. *J Hyg (Lond)* **83** (3): 547-55.
- 11 . Sellers, R. F. et D. E. Pedgley (1985).
Possible windborne spread to western Turkey of bluetongue virus in 1977 and of Akabane virus in 1979. *J Hyg (Lond)* **95** (1): 149-58.

- 12 . Sellers, R. F. et A. R. Maarouf (1991).
Possible introduction of epizootic hemorrhagic disease of deer virus (serotype 2) and bluetongue virus (serotype 11) into British Columbia in 1987 and 1988 by infected *Culicoides* carried on the wind. *Can J Vet Res* **55** (4): 367-70.
- 13 . Braverman, Y. et F. Checkik (1996).
Air streams and the introduction of animal diseases borne on *Culicoides* (Diptera, Ceratopogonidae) into Israel. *Rev Sci Tech* **15** (3): 1037-52.
- 14 . Mellor, P. S., J. Boorman et M. Baylis (2000).
Culicoides biting midges: their role as arbovirus vectors. *Annu Rev Entomol* **45**: 307-40.
- 15 . Ward, M. P. (1994).
The use of discriminant analysis in predicting the distribution of bluetongue virus in Queensland, Australia. *Vet Res Commun* **18** (1): 63-72.
- 16 . Ward, M. P. (1996).
Climatic factors associated with the infection of herds of cattle with bluetongue viruses. *Vet Res Commun* **20** (3): 273-83.
- 17 . Wittmann, E. J., P. S. Mellor et M. Baylis (2001).
Using climate data to map the potential distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in Europe. *Rev Sci Tech* **20** (3): 731-40.
- 18 . Baylis, M., H. Bouayoune, J. Touti et H. El Hasnaoui (1998).
Use of climatic data and satellite imagery to model the abundance of *Culicoides imicola*, the vector of African horse sickness virus, in Morocco. *Med Vet Entomol* **12** (3): 255-66.
- 19 . Baylis, M. et P. Rawlings (1998).
Modelling the distribution and abundance of *Culicoides imicola* in Morocco and Iberia using climatic data and satellite imagery. *Arch Virol Suppl* **14**: 137-53.
- 20 . Baylis, M., R. Meiswinkel et G. J. Venter (1999).
A preliminary attempt to use climate data and satellite imagery to model the abundance and distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in southern Africa. *J S Afr Vet Assoc* **70** (2): 80-9.
- 21 . Baylis, M., P. S. Mellor, E. J. Wittmann et D. J. Rogers (2001).
Prediction of areas around the Mediterranean at risk of bluetongue by modelling the distributions of its vector using satellite imaging. *Veterinary Record* **149**: 639-643.
- 22 . Tatem, A. J., M. Baylis, P. S. Mellor, B. V. Purse, R. Capela, I. Pena et D. J. Rogers (2003).
Prediction of bluetongue vector distribution in Europe and north Africa using satellite imagery. *Vet Microbiol* **97** (1-2): 13-29.
- 23 . Purse, B. V., A. J. Tatem, S. Caracappa, D. J. Rogers, P. S. Mellor, M. Baylis et A. Torina (2004).
Modelling the distributions of *Culicoides* bluetongue virus vectors in Sicily in relation to satellite-derived climate variables. *Med Vet Entomol* **18** (2): 90-101.
- 24 . Plazanet, D. (2002)
Application de a télédétection à l'épidémiologie de la fièvre catarrhale du mouton. CIRAD et Université Paul Sabatier (Toulouse III), Toulouse. 55 p.

- 25 . Gauthier, A. (2002).
La Corse. Une île-montagne au coeur de la Méditerranée. Paris, Delachaux et Niestlé. 320 p.
- 26 . Aas, G. et A. Riedmiller (2004).
Les arbres. Paris, Nathan. 256 p.
- 27 . Schönfelder, P. et I. Schönfelder (2004).
Flore du Bassin Méditerranéen. Paris, Vigot. 319 p.
- 28 . Girard, M. C. et C. Girard (1999).
Traitement des données de télédétection. Paris, Dunod. 529 p.
- 29 . Perez Correa, M. (2004)
Développement d'une méthode de cartographie de l'occupation du sol le long des cours d'eau à partir de données de télédétection. Cemagref, Montpellier.
- 30 . Bouyer, J., D. Hémon, S. Cordier, F. Derriennic, I. Stücker, B. Stengel et J. Clavel (1995).
Epidémiologie. Principes et méthodes quantitatives. Paris, Editions INSERM. 498 p.
- 31 . Biteau-Coroller, F. (2003)
Evaluation des performances du test de diagnostic cELISA utilisé pour le suivi de la fièvre catarrhale du mouton en France. Conséquences sur la prévalence réelle de la maladie et sur la stratégie d'échantillonnage. Université Joseph Fournier (Grenoble I); Université de Savoie (Annecy); Université René Descartes (Paris V); Université Henri Poincaré (Nancy I), Grenoble, Annecy, Paris et Nancy. 58 p.
- 32 . Rodhain, F. (2000).
Emergences et re-émergences des maladies infectieuses. Actes du Festival International de Géographie, Saint-Dié des Vosges, France, Géographie et Santé. Disponible sur le site : http://xxi.ac-reims.fr/fig-st-die/actes/actes_2000/rodhain/article.htm#haut.
- 33 . Garnier, M., V. Delamare, J. Delamare et T. Delamare (1999).
Dictionnaire des termes de médecine. Paris, Maloine. 973 p.
- 34 . Toma, B., B. Dufour, M. Sanaa, J.-J. Bénet, A. Shaw, F. Moutou et A. Louzã (2001).
Epidémiologie appliquée à la lutte collective contre les maladies animales transmissibles majeures. Paris, AEEMA. 696 p.
- 35 . Dupias, G. (1963).
Notice sommaire de la Carte de végétation de la France n° 80 et 81 : La Corse. Toulouse, Centre National de la Recherche Scientifique. 21 p.
- 36 . Rodhain, F. et C. Perez (1985).
Précis d'entomologie médicale et vétérinaire. Paris, Maloine. 458 p.

GLOSSAIRE

ACP :

Une analyse en composantes principales est une technique mathématique permettant de réduire un système complexe de variables corrélées entre elles en un plus petit nombre de dimensions. Chaque nouvelle dimension est constituée d'un assemblage des variables initiales. Une dimension est représentée par un axe qui porte le pourcentage de variance expliquée par les variables constituant l'axe.

Auto-corrélation spatiale :

L'auto-corrélation spatiale traduit le fait que deux variables proches dans l'espace ont des valeurs proches.

Emergentes :

Se dit de maladies qui ont été récemment découvertes, ou dont l'incidence et/ou la répartition géographique se sont brusquement accrues, ou qui ont atteint de nouvelles populations-hôtes [32].

Epizootie :

Se dit d'une maladie qui frappe simultanément un grand nombre d'animaux (équivalant à « épidémie » pour une maladie humaine) [33]. Se distingue d'une « enzootie », maladie animale sévissant à un taux relativement constant (équivalant à « endémique » pour une maladie humaine) [34].

Foyer :

Siège principal d'une maladie. Employé ici pour désigner les élevages ovins infectés ayant été confirmés par une analyse virologique.

Garrigue :

Voir maquis

Hématophage :

Se dit d'un insecte dont le cycle de développement nécessite un repas sanguin.

Isodata :

Méthode de classification non supervisée dite par « centres mobiles » qui calcule la moyenne des comptes numériques d'un groupe des classes réparties de façon homogène puis itérativement agrège les pixels en fonction des mesures de distance à ces groupes, les pixels étant affectés au groupe le plus proche.

Luminance :

Intensité énergétique émise par unité de surface apparente selon une direction donnée

Maquis (maquis, garrigue et pelouse) :

Paysage végétal le plus caractéristique de Corse, constitué d'arbustes et arbrisseaux xérophiles à feuillage persistant, en formation dense et difficilement pénétrable. Les espèces les plus fréquemment rencontrées sont l'arbousier et la bruyère en arbre, ainsi que le lentisque, le myrte, l'oléastre (olivier sauvage) et le ciste de Montpellier [35]. La dégradation du maquis par les feux et l'action de l'homme conduit à des formations plus basses par

l'élimination de l'arbousier et de la bruyère en arbre, appelées garrigue. Le terme ultime de cette dégradation est la pelouse, formation végétale basse ne dépassant pas 60 centimètres.

Matrice de confusion :

Permet d'estimer le pourcentage de pixels de l'image correctement classés en se référant à une « vérité-terrain » qui peut être une carte déjà existante ou des noyaux de validation.

Maximum de vraisemblance (classification par ~) :

Les classifications par maximum de vraisemblance se basent sur le calcul de la probabilité de chaque pixel d'être rattaché à une classe plutôt qu'à une autre. Une règle d'affectation permet de minimiser les risques d'erreur en utilisant au mieux les probabilités d'appartenance [28].

Méta-données

Terme utilisé pour désigner les informations décrivant les données : auteurs, date de recueil, traitements effectués...

Odds Ratio :

Dit aussi « rapport de côtes ». Est un indice de risque qui se définit comme le rapport des côtes chez les sujets exposés à un facteur sur celle des sujets non exposés à ce facteur. La côte est la valeur $p/(1 - p)$.

Pelouse

Voir maquis

Protection croisée :

Un sérotype A ne présentant pas de protection croisée avec un sérotype B induit la formation d'anticorps qui ne protègent pas contre le sérotype B.

Réfectance :

« Rapport de l'énergie réfléchi par un objet dans une bande de longueur d'onde à l'énergie reçue du soleil par cet objet et pour la même bande de longueur d'onde » [28].

Réservoir :

Décrit une « population, vertébrée ou invertébrée, assurant le maintien de l'agent infectieux dans la nature, notamment durant les périodes défavorables à la transmission, et remettant ensuite l'agent en circulation, permettant ainsi la reprise de la transmission » [36].

Scan 25® :

Produit de l'IGN correspondant au scannage des cartes à l'échelle 1:25 000.

Séparabilité de noyaux :

Traduit le niveau de confusion existant entre les noyaux de classes différentes.

Elle s'estime visuellement, grâce à un histogramme pluridimensionnel représentant pour chaque pixel, la probabilité d'appartenance à tel ou tel noyau ou en calculant pour les noyaux pris deux à deux le coefficient de Jeffries – Matusita.

Sérotype :

Catégorie dans laquelle on classe les virus et les bactéries selon leurs réactions en présence de sérums contenant des anticorps spécifiques [33]. Dans le cas de la fièvre catarrhale ovine, il existe 24 sérotypes différents.

Vecteur :

Définition au sens large : être vivant qui, à l'occasion de relations écologiques, acquiert un agent pathogène sur un hôte et le transmet ensuite à un autre hôte [34].

Définition au sens strict : arthropode hématophage qui assure la transmission biologique (ou mécanique) active d'un agent infectieux d'un vertébré à un autre vertébré [36]

Virémique :

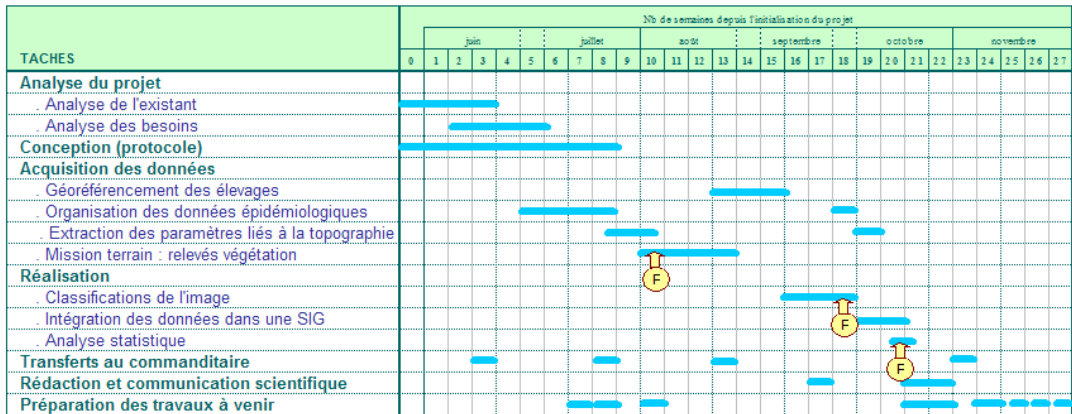
Décrit l'état d'un vertébré infecté chez qui le virus circule dans le sang.


ANNEXES

Annexe 1 : Modèles de répartition de *C. imicola*

Auteurs (année)	Variable expliquée (pays)	Variables explicatives (Variables testées)	Validité
[18; 19]	Abondance de <i>C. imicola</i> (Maroc)	<ul style="list-style-type: none"> ▪ Vitesse du vent (moyenne journalière de la vitesse minimale du vent pour le mois le moins venté de l'année) ▪ NDVI (moyenne annuelle du NDVI minimal) 	53 % de la variance expliquée
[20]	Abondance de <i>C. imicola</i> (Afrique du Sud)	<ul style="list-style-type: none"> ▪ Température de surface (minimale) (LST[#]) ▪ NDVI (minimal) 	67 % de la variance expliquée
[21]	Abondance de <i>C. imicola</i> (Maroc et péninsule ibérique)	<ul style="list-style-type: none"> ▪ NDVI (Phase du cycle annuel et trisannuel) ▪ MIR[#] (amplitude du cycle annuel) ▪ Altitude ▪ VPD[#] (phase du cycle annuel et trisannuel) ▪ Température de surface (phase du cycle annuel et minimum) 	Kappa = 0,9
[22]	Présence et abondance de <i>C. imicola</i> (Portugal)	<ul style="list-style-type: none"> ▪ NDVI (phase du cycle annuel et amplitude du cycle biannuel) ▪ MIR (moyen et amplitude du cycle biannuel) ▪ Température de surface (variance) ▪ Température de l'air (amplitude du cycle biannuel et phase du cycle trisannuel) 	Kappa = 0,88

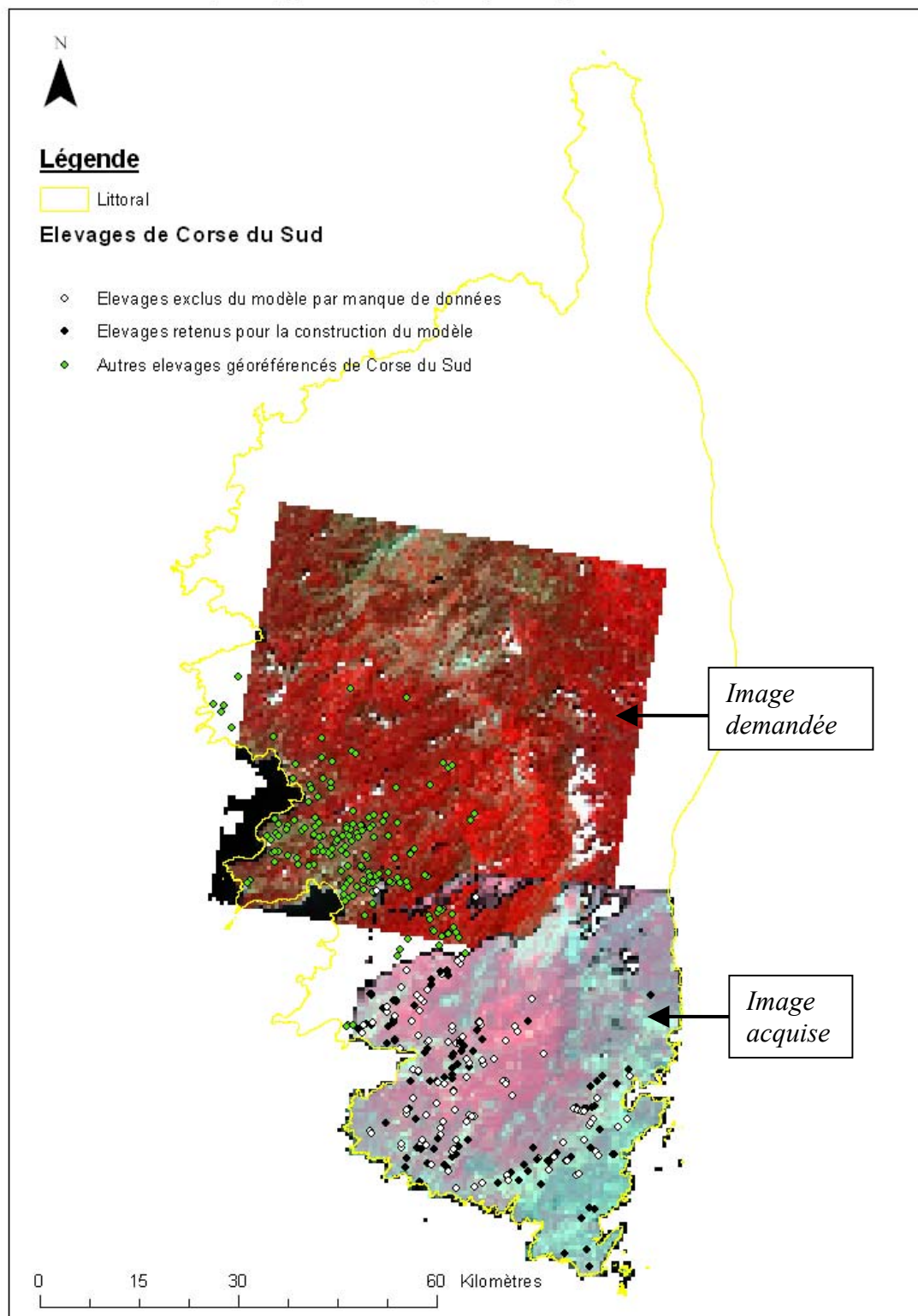
Annexe 2 : Chronogramme



 : Formations

Annexe 3 : Localisation des élevages de Corse du Sud par rapport à l'image satellitale SPOT

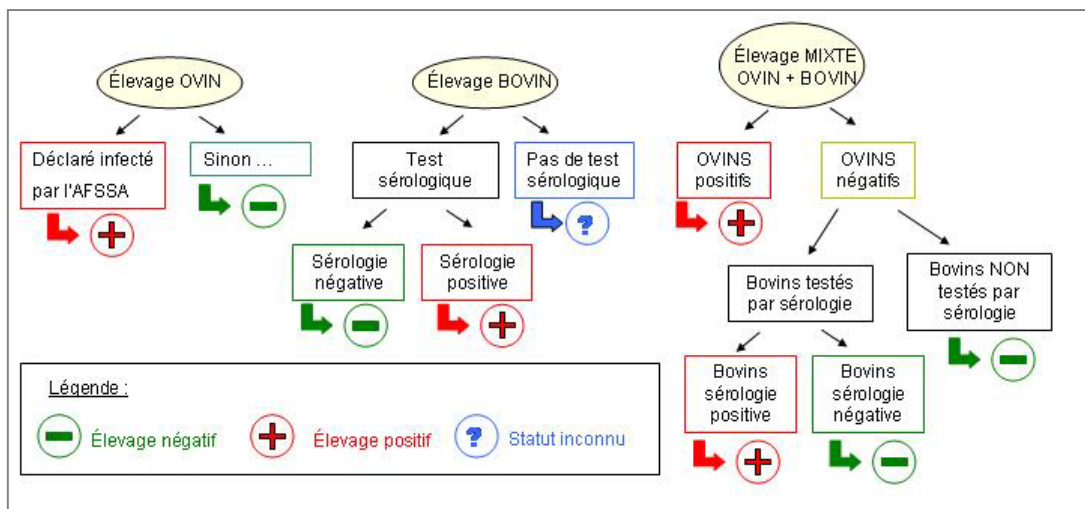
Position des élevages du sud de la Corse par rapport aux images Spot acquise et demandée



Donnée SPOT/ Programme ISIS, © CNES (2001), distribution Spot Image S.A.

Annexe 4 : Arbre de décision utilisable pour déterminer le statut sanitaire des élevages ovins, bovins et mixtes

Initialement, les élevages ovins, bovins et mixtes devaient être pris en compte dans l'étude. Seuls les élevages ovins ont finalement été conservés à cause des biais d'échantillonnage des élevages bovins. Si à l'avenir des données sur les bovins étaient recueillies sans biais d'échantillonnage, elles pourraient être incluses dans l'analyse. L'arbre de décision du statut sanitaire des élevages serait alors le suivant :



Légende :
 (-) Élevage négatif (+) Élevage positif (?) Statut inconnu

Annexe 5 : Fiches de relevés de végétation

Site n° :

Région :

Date : / / 2004

Heure : :

N° photo(s) terrain :
.....



Points GPS			
Nb			
N°	X (m) (Long)	Y(m) (Lat)	erreur (m)

Géomorphologie

thalweg	plaine	colline	versant	falaise	rive	vallée	crête

Occupation du Sol :

Non végétal :	<i>Bâti - urbain</i>	<i>Sol nu</i>	<i>Sable</i>	<i>Route</i>	<i>Roche</i>		
Forêt :	<i>Feuillus mélangés</i>	<i>Chêne</i>	<i>Châtaignier</i>	<i>Hêtre</i>	<i>Bouleau</i>	<i>Aune</i>	
	<i>Conifères mélangés</i>	<i>Pin</i>	<i>Sapin</i>				
Maquis	<i>lentisque</i>	<i>myrte</i>					
Garrigue	<i>lentisque</i>	<i>myrte</i>					
Pelouse	<i>lentisque</i>	<i>myrte</i>					
Culture :	<i>Verger</i>	<i>Vignoble</i>	<i>Oliveraie</i>	<i>Céréales</i>	<i>Prairie améliorée</i>		
Aquatique :	<i>Marais</i>	<i>Ripisylve</i>	<i>Rivière</i>				

Annexe 6 : Rapport de mission - Relevés de végétation effectués en Corse du Sud du 13 au 21 août 2004

Rapport de Mission en Corse du Sud du 13 au 21 août 2004

Mission n° 30 04 04 376

Nom : **Hélène Guis**



Fonction : Allocataire de recherche

Programme de rattachement : CIRAD EMVT ECONAP

Lieu de la mission : Corse du Sud

Dates : 13/08/2004 – 21/08/2004

Objectif :

- Effectuer des relevés de végétation en Corse du Sud pour réaliser une classification supervisée d'une image satellite SPOT acquise le 06/06/2001 ;
- Prendre contact avec la DSV de Corse du Sud ;
- Se rendre sur deux élevages ayant présenté des cas lors des précédentes épizooties de fièvre catarrhale ovine.

Calendrier :

- Samedi 14/08/04 : Arrivée à Propriano. Relevés de végétation dans la région située entre Propriano et Aullène.
- Dimanche 15/08/04 : Relevés de végétation dans la région située entre Aullène, Solenzara et Conca.
- Lundi 16/08/04 : Relevés de végétation dans les région de Porto Vecchio puis de Sartène. Rendez-vous avec M. G. Bousquet, Chef des Techniciens Vétérinaires, Antenne de Sartène.
- Mardi 17/08/04 : Relevés de végétation dans le sud-ouest : Figari, Monaccia d'Aullène, Sotta, Bonifacio.
- Mercredi 18/08/04 : Relevés de végétation dans la région située à l'ouest de Porto Vecchio. Rendez-vous avec M. Billandon, berger du troupeau de M. Marchetti (Bala), visite de l'élevage.
- Jeudi 19/08/04 : Relevés de végétation dans la région située autour de Propriano
- Vendredi 20/08/04 : Rendez-vous avec M. Cianfarani, éleveur à Pratazone, visite de la bergerie et des pâtures. Relevés de végétation dans la région située alentour. Rendez-vous avec M. J. Parodi, Adjoint Santé Animale, DSV D'Ajaccio. Départ d'Ajaccio.
- Samedi 21/08/04 : Retour Montpellier

Matériel :

- GPS, appareil photo numérique et ordinateur portable personnels ;
- Cartes IGN, Flores, Fiches de relevés de végétation ;
- Location d'une voiture sur place.

Bilan de la mission :

1. Relevés de végétation :

Une centaine de relevés de végétation (> 200 points GPS, > 200 photos numériques) ont été effectués (cf carte ci-jointe). Les sites de relevés ont été choisis pour leur homogénéité et leur taille suffisante. Ils ont été repérés in situ sur le terrain le long des routes ou bien au préalable sur les cartes IGN (zone marécageuse...) ou d'après une image obtenue en réalisant une ACP sur l'image satellite (l'ACP a permis de mettre en évidence des zones qui sont à la fois homogènes et bien distinctes entre-elles).

Ces relevés auront été effectués à la même saison que celle où l'image a été acquise (06/06/2001), période propice à la multiplication des Culicoides.

Les relevés effectués vont permettre de classer l'image satellite du Sud de la Corse du 06/06/2001 afin de rechercher les paysages à risques en étudiant les corrélations entre statut des élevages (infectés/indemnes) et les types de végétation situés autour des élevages.

2. Contact avec la DSV de Corse du Sud

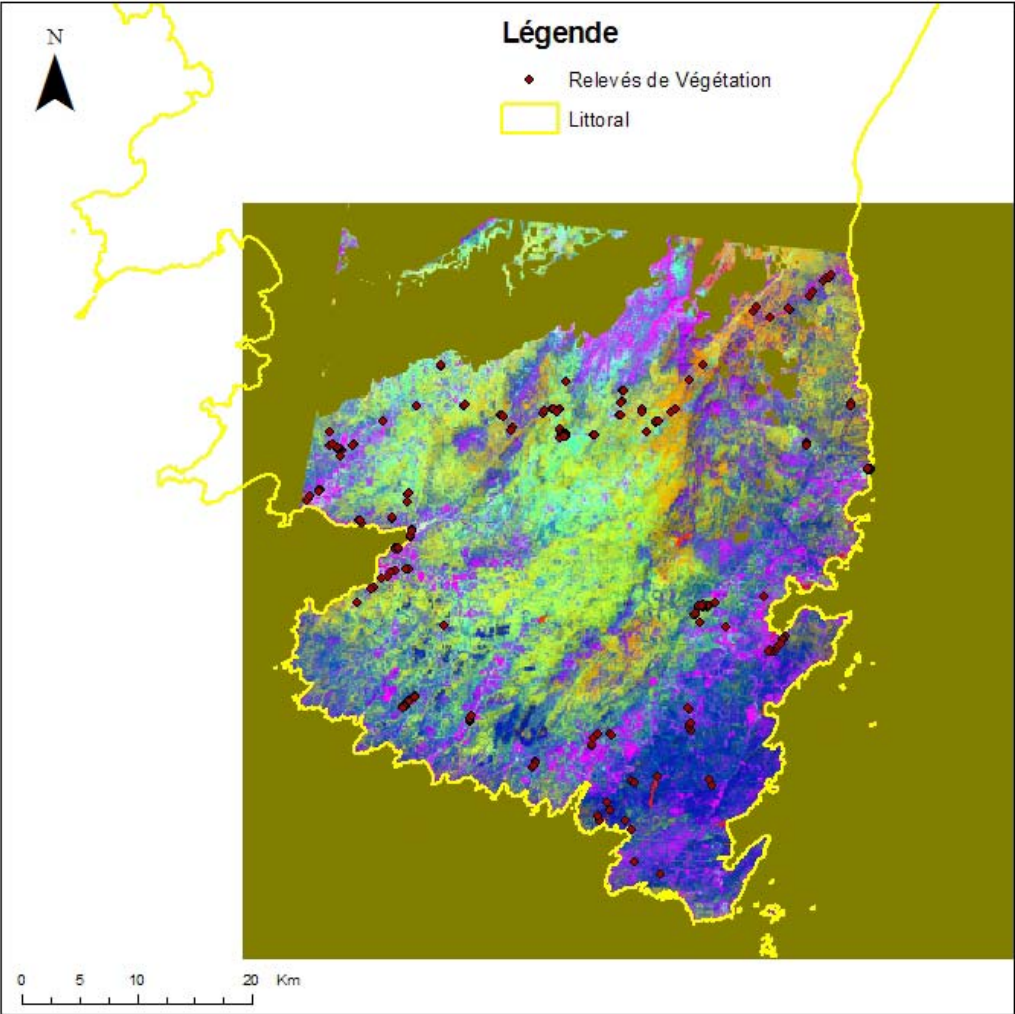
La mission a été l'occasion de prendre contact et d'informer la DSV du travail en cours. Les discussions ont porté sur le transfert des données entre le CIRAD et la DSV, la politique de vaccination, l'arrivée potentielle du sérotype 16, leurs perceptions des facteurs de risque, et la perception et l'attitude des éleveurs face à la FCO. M. Bousquet nous a mis en relation avec deux éleveurs qui ont accepté de nous recevoir. Un troisième éleveur (M. Andreani, Figari) n'a pu être joint.

3. Visite de deux élevages ayant été infectés par la FCO

Deux élevages infectés ont été visités. Il aurait été intéressant de visiter également des élevages situés à proximité mais non infectés ou peu infectés (notamment à Pratazone où les autres élevages ont eu peu de cas alors que les pâtures sont situées tout près de celles de M. Cianfarani). En visitant deux élevages, l'objectif n'était pas de formuler des hypothèses concernant les facteurs de risque mais plutôt d'être confronté sur le terrain à des zones où *Culicoides imicola* est présent et la FCO a circulé, de comprendre la conduite d'élevage en place, et de visiter les pâtures où sont les troupeaux. Cela permettra aussi de discuter avec G. Gerbier et T. Baldet qui ont également visité ces deux élevages lors d'une précédente mission. Remarque 1 : Les élevages présentaient tous deux des zones humides à proximité, notamment celui de M. Cianfarani, les différents ruisseaux dans les pâtures n'étant pas asséchés à la période de la mission (mi-août).

Remarque 2 : M. Billandon a fait part de sa volonté de construire un système d'écoulement reliant la bergerie à une fosse à fumier. Des piègeages réalisés auparavant par T. Baldet avaient permis de capturer de nombreux *C. imicola* à cet endroit. Il sera intéressant de voir l'évolution en termes de densité de *C. imicola* suite à la réalisation de ces travaux.

Relevés de végétation effectués en Corse du Sud



Annexe 7 : Attributs choisis et retenus pour la classification orientée objet par plus proches voisins

En tout, 66 attributs ont été choisis pour réaliser la classification orientée objet par plus proche voisins. La meilleure classification est celle présentant 7 dimensions (donc 7 attributs sont retenus pour réaliser la classification).

1. Ensemble des 66 attributs choisis pour réaliser la classification orientée objet par plus proches voisins

Pour chacune des 4 bandes que comporte l'image SPOT, les attributs suivants ont été inclus :

- attributs concernant les valeurs radiométriques : moyenne, écart type, ratio, écart au pixels voisins, valeurs minimale et maximale des pixels,
- attributs de texture :
 - o GLCM (*Grey Level Cooccurrence Matrix*) : la dissimilitude, le contraste, l'homogénéité, entropie, second moment (ou moment inverse des différences), moyenne (ou énergie)
 - o GLDV (*Grey Level Difference Vector*) : moyenne et second moment

A ces attributs s'ajoutent ceux calculés pour toutes les bandes confondues : l'aire, « brightness », et les attributs de texture présentés ci-dessus.

2. Attributs retenus pour la classification orientée objet par plus proches voisins :

(La meilleure classification présente 7 dimensions)

- 1 - Valeur minimale des pixels (bande 1)
- 2 - GLCM Homogénéité (bande 3)
- 3 - Aire
- 4 - Ratio (bande 4)
- 5 - GLCM Second moment
- 6 - Valeur maximale des pixels (bande 3)
- 7- Valeur maximale des pixels (bande 2)

1. Amélioration des classifications pixel à pixel

Les améliorations de la classification pixel à pixel ont nécessité un grand nombre d'étapes pour passer de 30 types de noyaux initiaux à 8 classes de végétation finales. Les différentes approches mises en œuvre ou testées sont :

- différentes classes de maquis-garrigue-pelouse ont été testées selon les espèces végétales qu'elles comportaient ou leur valeurs dans le PIR (+/- lentisque, +/- myrte, +/- pins laricio, +/- ciste de Montpellier, +/- chlorophyllien),
- redéfinition des ROI de certaines classes (par exemple la classe pelouse comportant roches),
- fusion des noyaux d'apprentissages selon leur séparabilité,
- regroupement de classes différentes selon leur ressemblance sémantique (par exemple : eau de mer, eau trouble, et eau limpide),
- classification sur les 4 canaux initiaux de l'image et/ou combinés avec 3 néo-canaux de l'ACP portant le plus de variance et/ou un canal constitué du NDVI,
- classification à partir de l'ensemble des noyaux (d'apprentissage et de validation) dans l'optique d'une validation grâce à Corine Land Cover,
- masquage des classes qui se séparaient bien des autres pour relancer des classifications uniquement sur les classes restantes,
- test de variation du seuil de classification par maximum de vraisemblance,
- essai de classification par isodata avec 50 classes.

2. Amélioration des classifications orientées objet

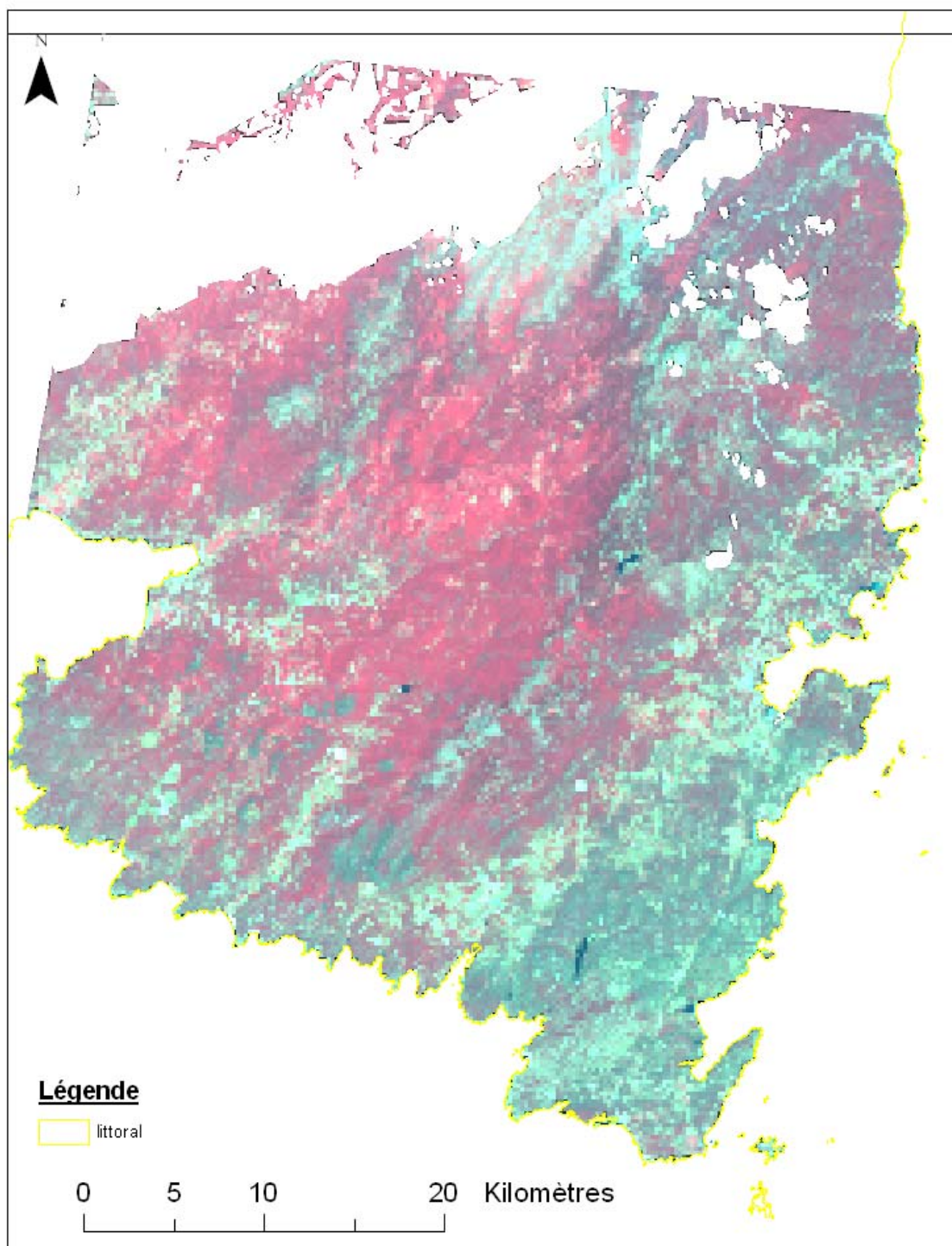
Les améliorations de la classification orientée objet ont été obtenues grâce aux étapes suivantes :

- plusieurs types de segmentation ont été testées jusqu'à obtention d'une segmentation assez fine,
- différents seuils de classification ont été testés,
- plusieurs dimensions (de 5 à 12) ont été testées. Finalement c'est avec une dimension égale à 7 que la meilleure classification a été obtenue.
- enfin, les classes noyaux d'apprentissage ont été fusionnés et les classes combinées pour obtenir 9 classes finales.

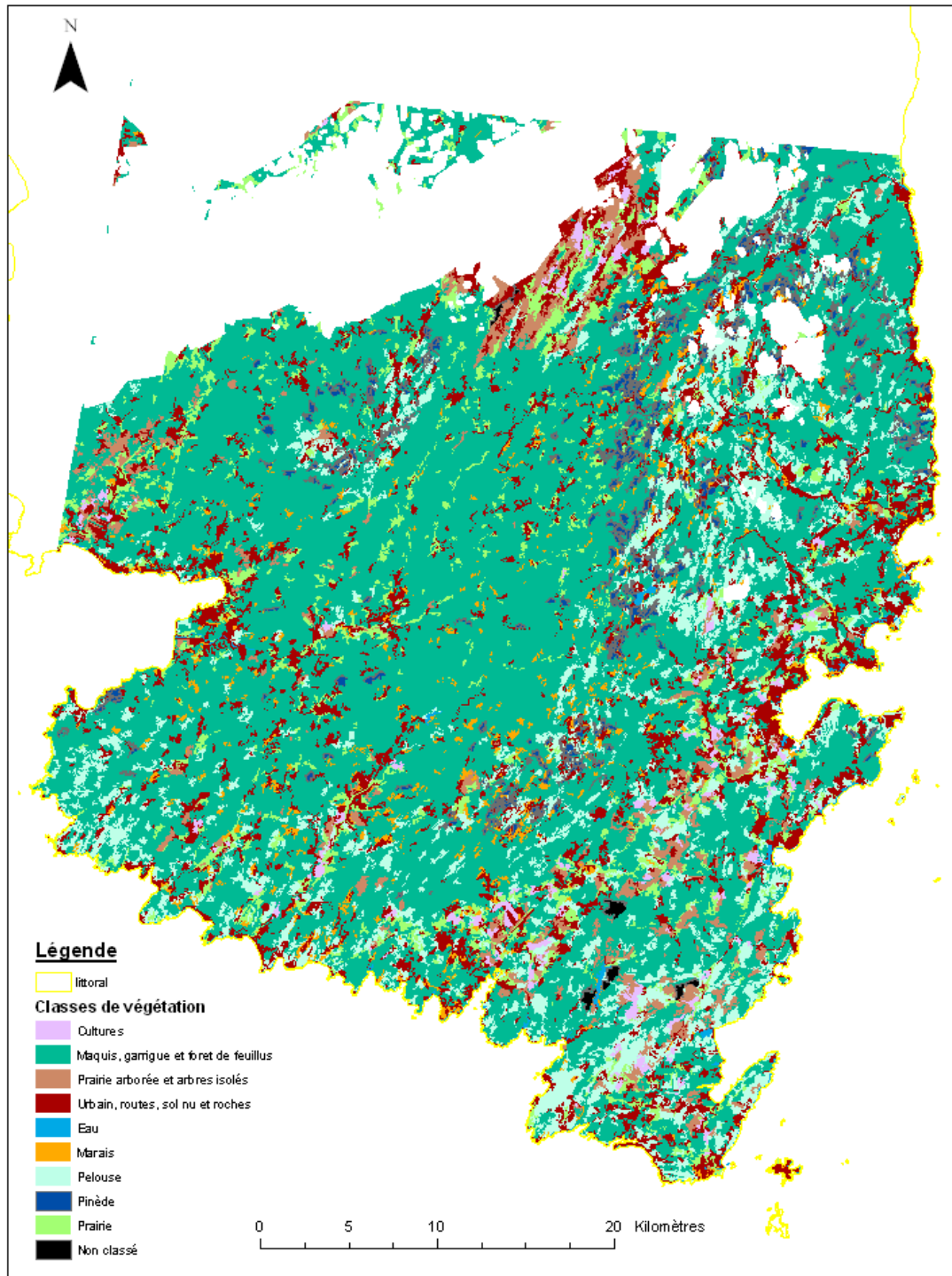
Annexe 9 : Classifications finales

L'image initiale et classifiée selon les deux méthodes sont présentées dans leur ensemble puis une zone est détaillée. Pour la classification pixel à pixel, la vue générale n'a pas été présentée car est très modifiée par rapport à l'image originale, sans doute à cause du redimensionnement de l'image lors de son exportation.

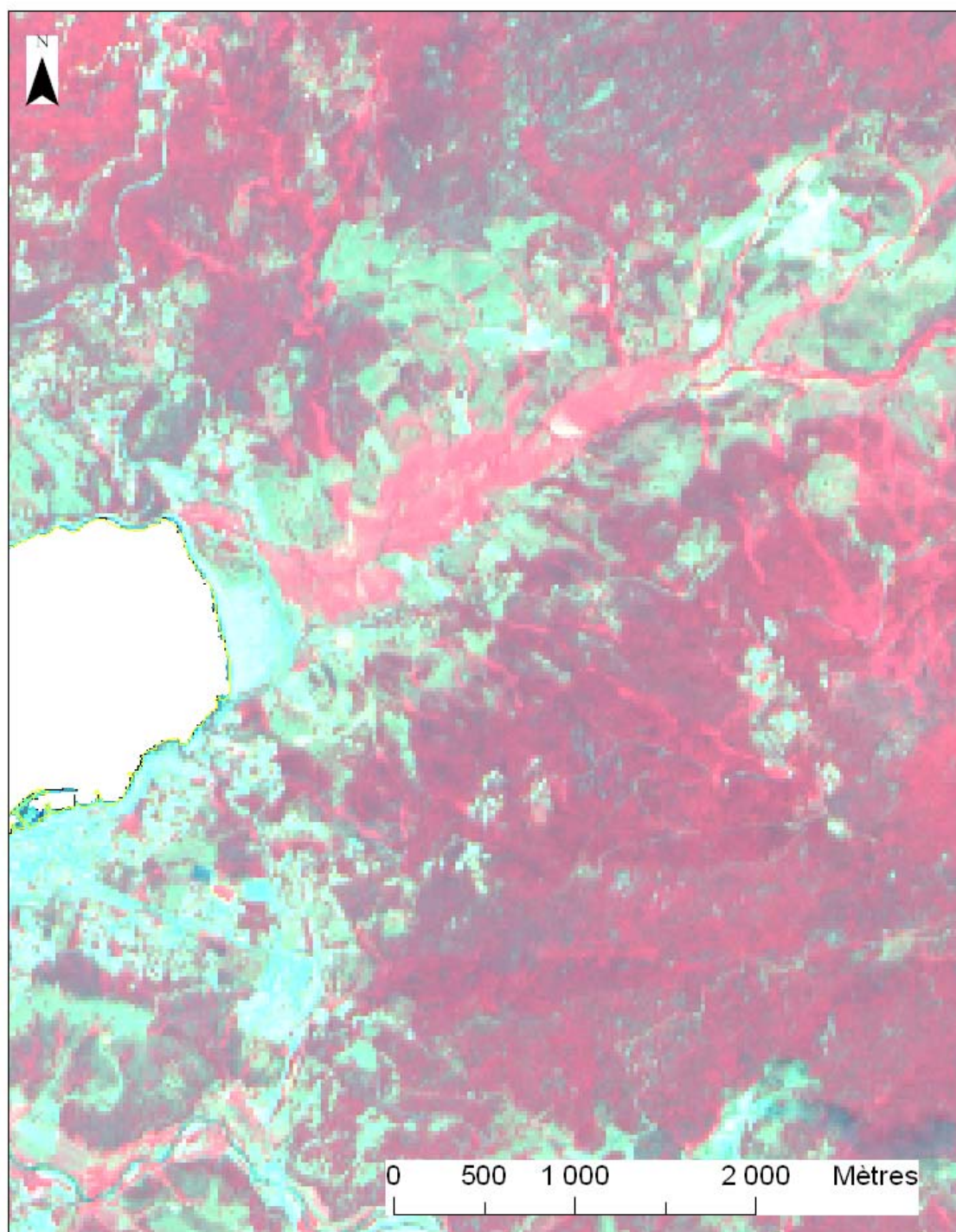
Composition colorée de l'image SPOT du sud de la Corse



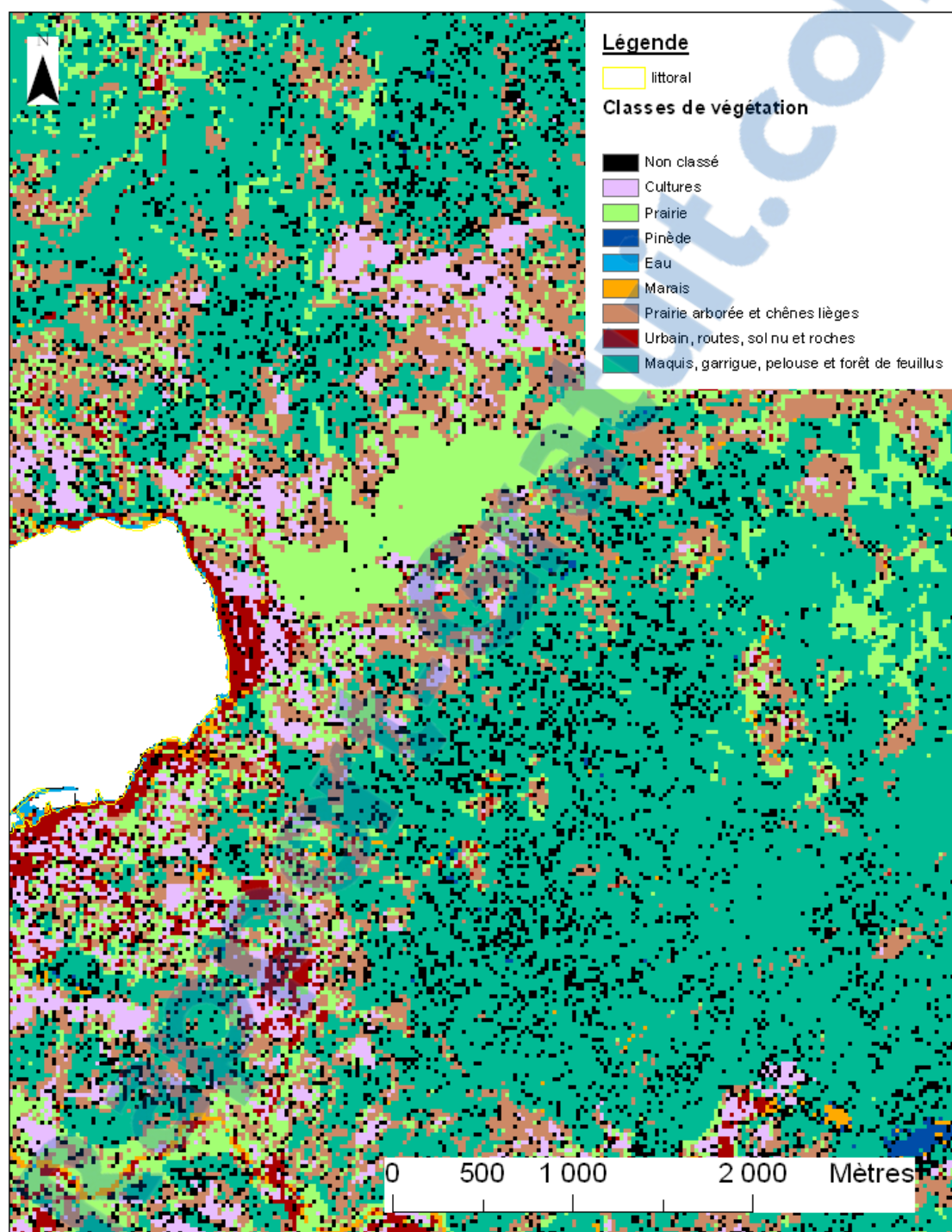
Classification orientée objet de l'image SPOT du sud de la Corse



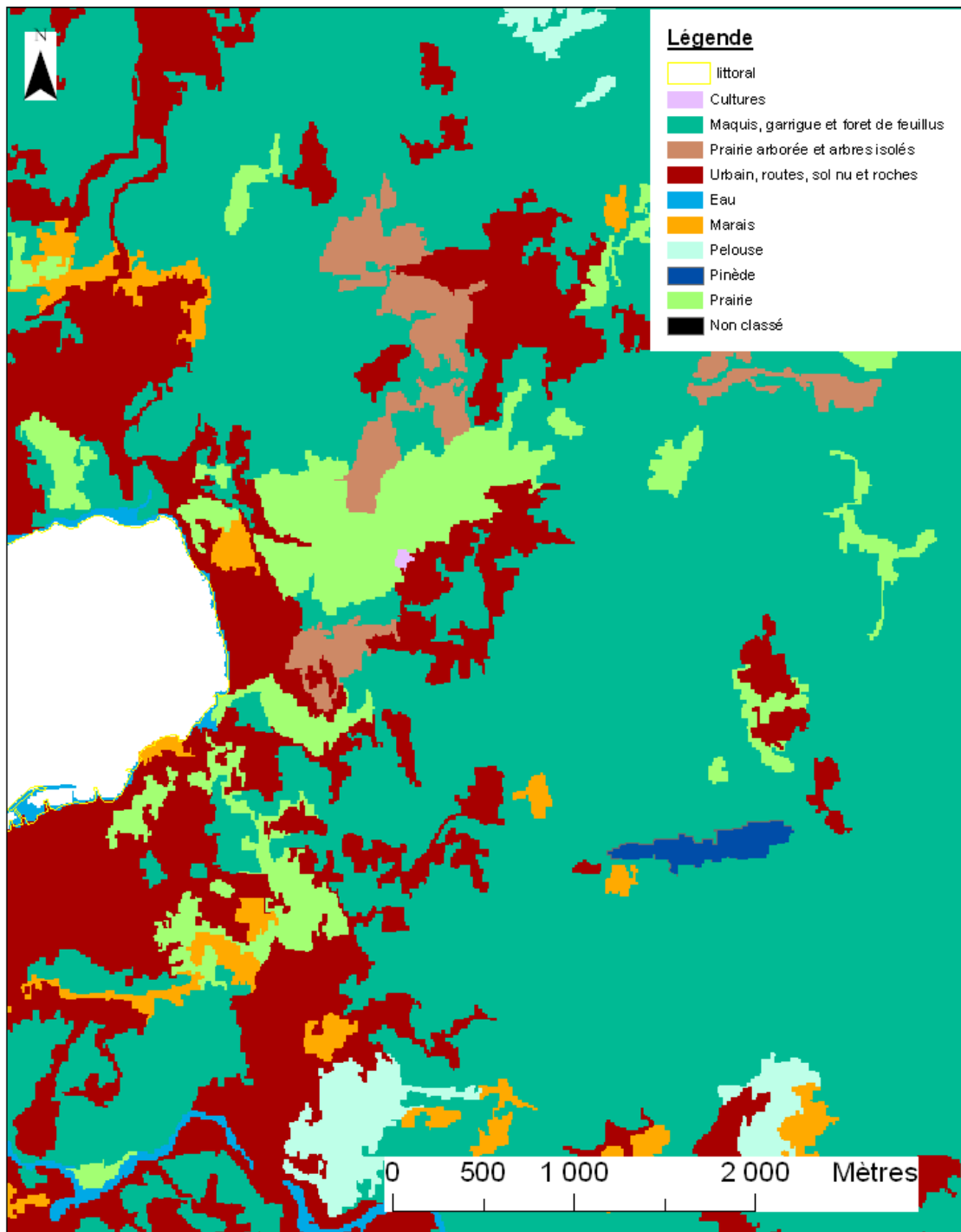
Détail d'une composition colorée de l'image SPOT du sud de la Corse :
Baie de Propriano



Détail de la classification pixel à pixel de l'image SPOT du sud de la Corse :
Baie de Propriano



Détail de la classification orientée objet de l'image SPOT du sud de la Corse :
Baie de Propriano



Annexe 10 : Matrices de confusion des classifications pixel à pixel et orientée objet

1. Matrice de confusion de la classification pixel à pixel :

Classes	Noyaux de validation								Total
	pinède	bâti_sol nu	marais	eau	prairie	culture	prairie arborée- chênes lièges	maquis-garrigue- pelouses	
non classé	10.81	13.04	13.48	0.50	6.71	7.98	12.32	13.46	12.03
pinède	74.29	0.00	0.94	0.18	0.00	0.00	0.00	0.56	3.62
bâti_sol nu	0.00	67.72	0.27	0.82	33.54	7.29	0.00	0.01	16.46
marais	3.18	0.28	39.08	5.33	0.00	0.00	0.00	0.71	1.20
eau	0.00	0.02	10.51	92.76	0.18	0.00	0.00	0.00	3.08
prairie	0.00	8.31	1.35	0.05	12.91	1.02	6.87	3.70	4.84
culture	0.00	10.24	0.54	0.00	45.52	83.61	16.36	6.55	14.10
prairie arborée- chênes lièges	0.00	0.30	1.89	0.00	1.11	0.02	62.42	0.50	0.90
maquis-garrigue- pelouses	11.73	0.09	31.94	0.36	0.04	0.07	2.02	74.51	43.77
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Précision totale : 69.37 %

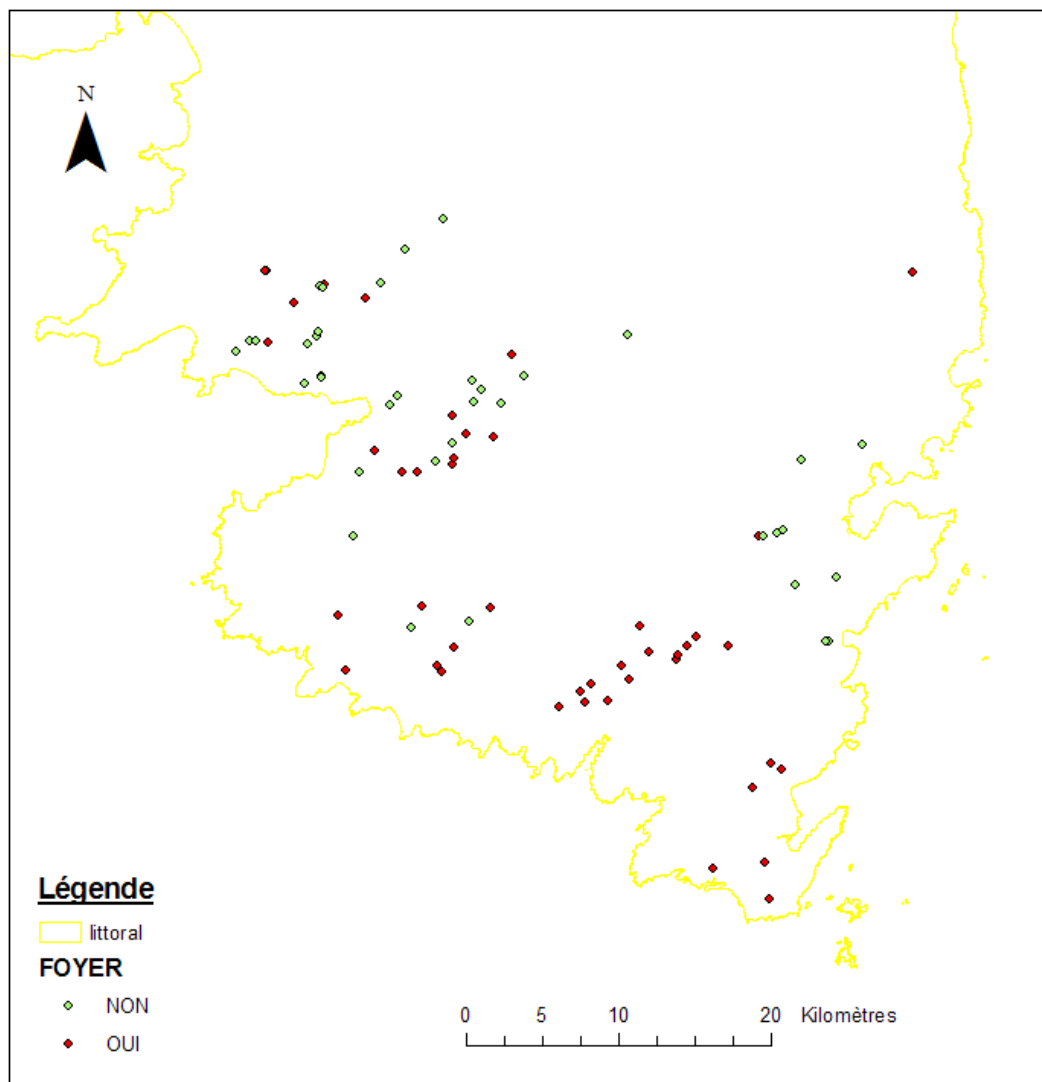
Coefficient de Kappa : 0.5628

2. Matrice de confusion de la classification orientée objet (en %) :

Classification	Noyaux de validation									Total
	pinède	bâti-sol nu	prairie	prairie arborée - arbres isolés	cultures	marais	pelouse	maquis-garrigue- forêt de feuillus	eau	
pinède	92.52	0,00	0,00	0,00	0,00	0,00	0,00	4,58	0,00	5,98
bâti-sol nu	0,00	65.97	8,03	7,17	30,73	0,00	0,00	0,00	0,43	31,45
prairie	0,00	13,46	73,04	0,00	0,00	0,00	0,00	3,08	0,00	9,09
prairie arborée - arbres isolés	0,00	3,19	0,00	75,43	1,10	0,00	33,21	0,00	0,00	5,94
cultures	0,00	10,02	0,00	10,59	68,16	0,00	0,00	0,00	0,00	6,07
marais	0,00	0,00	0,00	0,00	0,00	5,37	0,00	0,00	0,00	0,04
pelouse	0,00	3,88	13,67	6,81	0,00	0,00	66,79	6,44	0,00	8,80
maquis-garrigue-forêt de feuillus	7,48	2,57	5,25	0,00	0,00	87,79	0,00	85,90	0,00	27,16
eau	0,00	0,91	0,00	0,00	0,00	6,84	0,00	0,00	99,57	5,46
non classé	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Total	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Précision totale : 74,80 %

Statut sanitaire des élevages ovins de Corse du Sud



Annexe 12 : Analyse descriptive et corrélations entre les variables

1. Analyse descriptive des variables

Variables Quantitatives

	Elevages négatifs moyenne (écart -type)	Elevages positifs moyenne (écart -type)	Probabilité
X	555647.102 (13108.711)	558057.677 (11208.061)	0.357
Y	4152184.369 (7890.004)	4143628.961 (11760.266)	0.000
NDVI MOYEN	- 0.021 (0.119)	- 0.085 (0.099)	0.008
NDVI MINIMUM	- 0.394 (0.074)	- 0.411 (0.054)	0.209
NDVI MAXIMUM	0.342 (0.077)	0.302 (0.097)	0.035
NDVI Amplitude	0.735 (0.065)	0.714 (0.098)	0.220
ENSOLEILLEMENT MOYEN	0.436 (0.214)	0.441 (0.184)	0.909
PENTE MOYENNE	3.210 (2.934)	4.068 (3.238)	0.193
ALTITUDE MOYENNE	145.786 (169.580)	97.128 (108.351)	0.116
Classification pixel à pixel			
NON CLASSE	13.582 (5.274)	16.880 (7.746)	0.020
CULTURE	17.498 (18.486)	24.487 (17.038)	0.068
PRAIRIE	19.905 (12.855)	14.280 (12.356)	0.039
PINEDE	0.194 (0.591)	0.245 (1.216)	0.800
EAU	0.071 (0.374)	0.005 (0.015)	0.263
MARAIS	0.734 (1.537)	0.865 (2.377)	0.755
PRAIRIE ARBOREE	16.786 (8.070)	14.245 (8.827)	0.160
CHENES LIEGES			
BATI-SOL NU	1.044 (1.193)	1.427 (1.775)	0.232
MAQUIS- GARRIGUE- PELOUSE-FORET	30.185 (17.826)	27.565 (14.881)	0.457
Classification orientée objet			
BATI-SOL NU	16.920 (12.314)	20.170 (16.916)	0.267
CULTURES	7.262 (13.784)	4.830 (9.696)	0.344
EAU	0.071 (0.463)	0.106 (0.729)	0.786
MAQUIS- GARRIGUE-FORET	52.333 (25.591)	40.021 (21.268)	0.016
MARAIS	0.595 (2.198)	2.085 (3.983)	0.030
PELOUSE	4.738 (7.402)	10.872 (14.105)	0.011
PINEDE	0.000	0.000	Insufficient data
PRAIRARBOI	6.000 (9.435)	8.064 (12.216)	0.372
PRAIRIE	12.310 (11.736)	13.766 (12.137)	0.567
NON CLASSE	0.000	0.000	Insufficient data

Variables Qualitatives

Frequencies

MIXITE :

	Test statistic	Value	df	Prob
Fisher exact test (two-tail)				0.012

Frequencies

ORIENTATION MAJORITAIRE (rows) by FOYER (columns)

	0	1	Total
0	26	27	53
1	16	20	36
Total	42	47	89

	Test statistic	Value	df	Prob
Pearson Chi-square		0.183	1.000	0.669

2. Matrice de corrélations des variables deux à deux

(Pearson correlation matrix)

Les classes notées (1) sont issues de la classification orientée objet, les classes notées (2) sont issues de la classification pixel à pixel.

	BATI	GGRS	CULTURES	EAUTOT	MAQGARRFOR	MARAIS
BATI_GGRS (1)		1.000				
CULTURES (1)	-0.024		1.000			
EAUTOT (1)	0.174	0.039		1.000		
MAQGARRFOR (1)	-0.455	-0.532	-0.102		1.000	
MARAIS (1)	-0.019	-0.033	-0.067	-0.176		1.000
PELOUSE (1)	-0.125	-0.032	0.107	-0.306		0.031
PINEDE (1)	-0.172	-0.077	-0.025	0.228		0.093
PRAIRARBOI (1)	0.019	0.364	-0.072	-0.408		-0.135
PRAIRIE (1)	-0.141	-0.053	-0.158	-0.281		0.224
UNCLASSIFI (1)	-0.120	0.195	-0.014	-0.169		-0.042
X	-0.037	0.323	-0.044	-0.240		-0.028
Y	-0.047	-0.229	-0.111	0.463		-0.215
MEANSOLEIL	0.054	-0.139	0.051	0.053		-0.066
PENTEMOY	0.072	-0.146	0.075	-0.082		0.030
ALTMOY	-0.254	-0.311	-0.115	0.588		-0.132
UNCLASSENV (2)	-0.068	-0.005	0.278	-0.253		0.036
CULTUREENV (2)	0.196	0.675	-0.028	-0.764		0.098
PRAIRIEENV (2)	0.212	-0.128	-0.065	0.194		-0.193
PINEDEENV (2)	-0.125	-0.058	-0.000	0.092		0.073
EAUENV (2)	0.082	0.208	0.796	-0.100		-0.030
MARAISENV (2)	-0.023	-0.032	0.235	0.033		0.062
PRAIRARBOCHL (2)	0.178	-0.235	-0.091	0.002		0.153
BATIGGRSENV (2)	0.170	0.142	0.393	-0.360		-0.099
MAQGARPELFOR (2)	-0.448	-0.511	-0.074	0.799		-0.045
NDVIMOY	-0.217	-0.517	-0.117	0.806		-0.097
NDVIMIN	-0.316	-0.446	-0.588	0.571		0.059
NDVIMAX	0.050	-0.326	-0.079	0.396		-0.101
NDVIRANG	0.316	0.048	0.409	-0.095		-0.145

	PELOUSE	PINEDE	PRAIRARBOI	PRAIRIE	UNCLASSIFI
PELOUSE	1.000				
PINEDE	0.039	1.000			
PRAIRARBOI	-0.169	-0.105	1.000		
PRAIRIE	-0.129	-0.173	-0.151	1.000	
UNCLASSIFI	-0.062	-0.016	0.480	-0.037	1.000
X	0.285	-0.148	0.035	-0.076	0.137
Y	-0.488	0.043	0.021	-0.110	-0.152
MEANSOLEIL	0.077	-0.060	-0.084	-0.037	0.011
PENTEMOY	0.175	-0.028	-0.093	0.085	0.059
ALTMOY	-0.140	0.123	-0.194	-0.249	-0.056
UNCLASSENV	0.559	0.018	0.052	-0.070	0.123
CULTUREENV	0.247	-0.143	0.317	0.151	0.182
PRAIRIEENV	-0.450	-0.091	0.050	-0.038	-0.094
PINEDEENV	0.091	-0.040	-0.091	-0.028	-0.023
EAUENV	0.017	-0.026	-0.033	-0.132	-0.015
MARAISENV	0.151	-0.039	-0.127	-0.102	0.004
PRAIRARBOCHL	-0.347	-0.066	0.020	0.306	-0.170
BATIGGRSENV	0.210	-0.104	0.206	0.001	0.338
MAQGARPELFOR	-0.018	0.269	-0.412	-0.240	-0.118
NDVIMOY	-0.411	0.141	-0.320	-0.168	-0.199
NDVIMIN	-0.206	0.287	-0.179	0.022	-0.101
NDVIMAX	-0.343	-0.058	-0.295	0.081	-0.361
NDVIRANG	-0.161	-0.299	-0.138	0.068	-0.265

	X	Y	MEANSOLEIL	PENTEMOY	ALTMOY
X	1.000				
Y	-0.513	1.000			
MEANSOLEIL	0.069	-0.001	1.000		
PENTEMOY	0.107	-0.168	0.450	1.000	
ALTMOY	-0.093	0.447	0.392	-0.006	1.000
UNCLASSE NVI	0.401	-0.557	-0.014	0.194	-0.248
CULTURE NVI	0.426	-0.633	-0.091	0.090	-0.552
PRAIRIE NVI	-0.429	0.657	-0.007	-0.207	0.279
PINEDEE NVI	0.153	0.141	-0.097	-0.109	0.162
EAU NVI	0.046	-0.130	0.015	-0.006	-0.123
MARAISE NVI	0.139	0.004	-0.086	-0.047	0.043
PRAIRARBOCHL	-0.391	0.325	0.016	-0.060	-0.138
BATIGGRRSENV	0.258	-0.263	0.092	0.209	-0.206
MAQGARPELFOR	-0.141	0.252	0.108	0.006	0.555
NDVIMOY	-0.448	0.709	0.075	-0.178	0.651
NDVIMIN	-0.290	0.448	0.066	-0.161	0.505
NDVIMAX	-0.437	0.608	0.097	-0.086	0.388
NDVIRANG	-0.191	0.219	0.035	0.051	-0.056

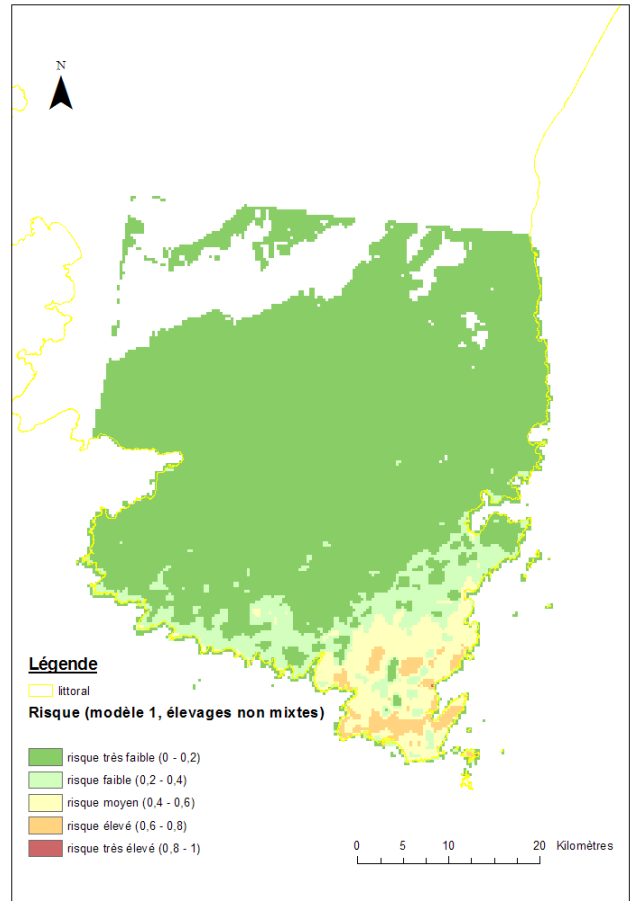
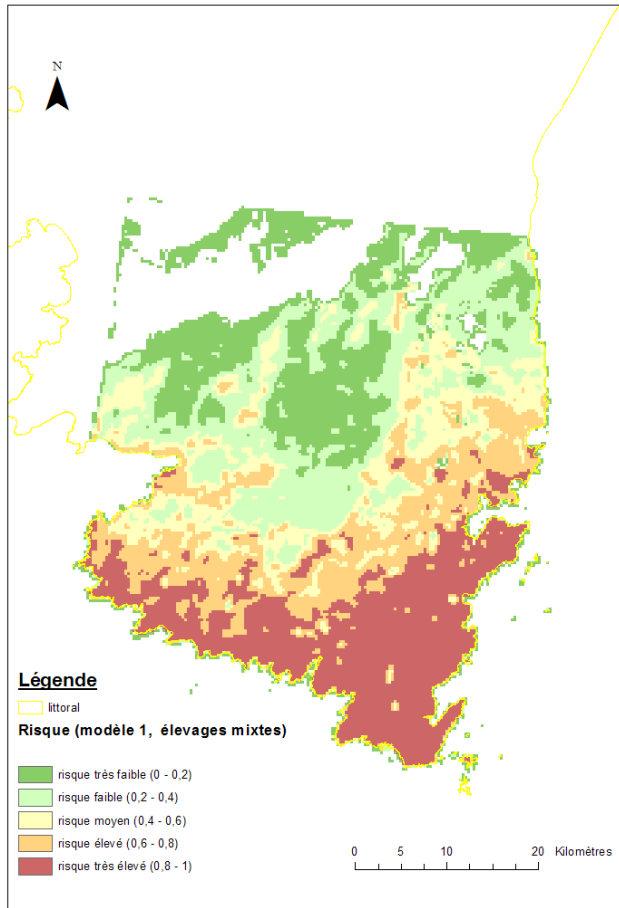
	UNCLASSE NVI	CULTUREE NVI	PRAIRIEE NVI	PINEDEE NVI	EAU NVI
UNCLASSE NVI	1.000				
CULTUREE NVI	0.248	1.000			
PRAIRIEE NVI	-0.577	-0.485	1.000		
PINEDEE NVI	0.186	-0.111	-0.101	1.000	
EAU NVI	0.240	0.082	-0.085	0.032	1.000
MARAISE NVI	0.361	-0.090	-0.163	0.844	0.262
PRAIRARBOCHL	-0.437	-0.260	0.287	-0.126	-0.136
BATIGGRRSENV	0.448	0.174	-0.127	0.011	0.216
MAQGARPELFOR	-0.120	-0.662	-0.134	0.029	-0.128
NDVIMOY	-0.602	-0.849	0.553	-0.013	-0.169
NDVIMIN	-0.410	-0.546	0.262	0.080	-0.573
NDVIMAX	-0.682	-0.482	0.570	-0.071	-0.121
NDVIRANG	-0.318	-0.017	0.341	-0.143	0.356

	MARAISE NVI	PRAIRARBOCHL	BATIGGRRSENV	MAQGARPELFOR
MARAISE NVI	1.000			
PRAIRARBOCHL	-0.196	1.000		
BATIGGRRSENV	0.171	-0.284	1.000	
MAQGARPELFOR	-0.015	-0.207	-0.257	1.000
NDVIMOY	-0.142	0.217	-0.425	0.681
NDVIMIN	-0.167	0.224	-0.552	0.520
NDVIMAX	-0.228	0.385	-0.350	0.230
NDVIRANG	-0.090	0.193	0.116	-0.213


	NDVIMOY	NDVIMIN	NDVIMAX	NDVIRANG
NDVIMOY	1.000			
NDVIMIN	0.663	1.000		
NDVIMAX	0.690	0.387	1.000	
NDVIRANG	0.117	-0.454	0.645	1.000

Annexe 13 : Essais de cartes d'aléa de fièvre catarrhale ovine


*Essais de carte d'aléa de fièvre catarrhale ovine
pour les élevages mixtes (à gauche) et non mixtes (à droite)*



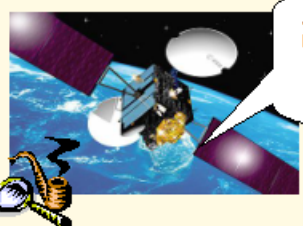
Annexe 14 : Poster de vulgarisation scientifique réalisé pour la fête de la science 2004



Des satellites à la recherche d'un moucheron...




Hélène Guis



Bonjour
Je suis Sherlock, le satellite SPOT. En tant que détective privé, je mène une enquête pour rechercher un dangereux moucheron tueur de moutons.


Aide-moi à le retrouver, il ressemble à ça :

RECHERCHÉ (mort ou vif)




Nom : Culicoides
Prénom : imicola
Recherché pour : Meurtre de moutons

Mais comment un si petit moucheron peut-il tuer un gros mouton ?




Le moucheron imicola attaque ses victimes à la tombée de la nuit pour boire un peu de leur sang. Il ne les tue pas parce qu'il boit leur sang mais parce qu'il leur transmet un virus qui les rend très malades et même les tue parfois.

Virus de la **BLUETONGUE**, de la « langue bleue » :
(fièvre catarrhale ovine)




Et moi Spot, vieux détective, comment je peux faire pour retrouver un si petit moucheron ?

Je suis un satellite qui n'est plus tout jeune, et je ne vois plus très bien : je ne distingue que les objets de ... 10 mètres




alors qu'imicola mesure 2 millimètres !!



X 5 000

Alors comment faire... As-tu une idée ?


Sais-moi !



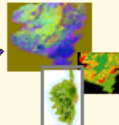
Ainsi, pour trouver l'habitat du moucheron imicola :

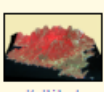
Il faut regrouper toutes sortes d'informations :

des images satellites
(vues par le détective SPOT)




des cartes de végétation







l'altitude
(modèle numérique de terrain)

des cartes de piéages de mouchérons




des paramètres climatiques (soleil, pluie, vent ...)





la localisation des élevages de moutons



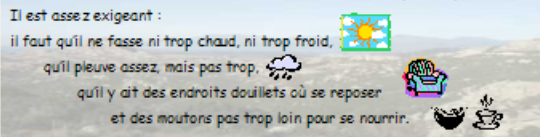
Carte de RISQUE de présence d'imicola


Merci Spot!

J'ai une solution : cherchons les endroits où il habite !

Au lieu de chercher directement le moucheron, je vais chercher son habitat (là où il habite). Il habite dehors, dans la végétation, mais pas n'importe où ! Il est assez exigeant :

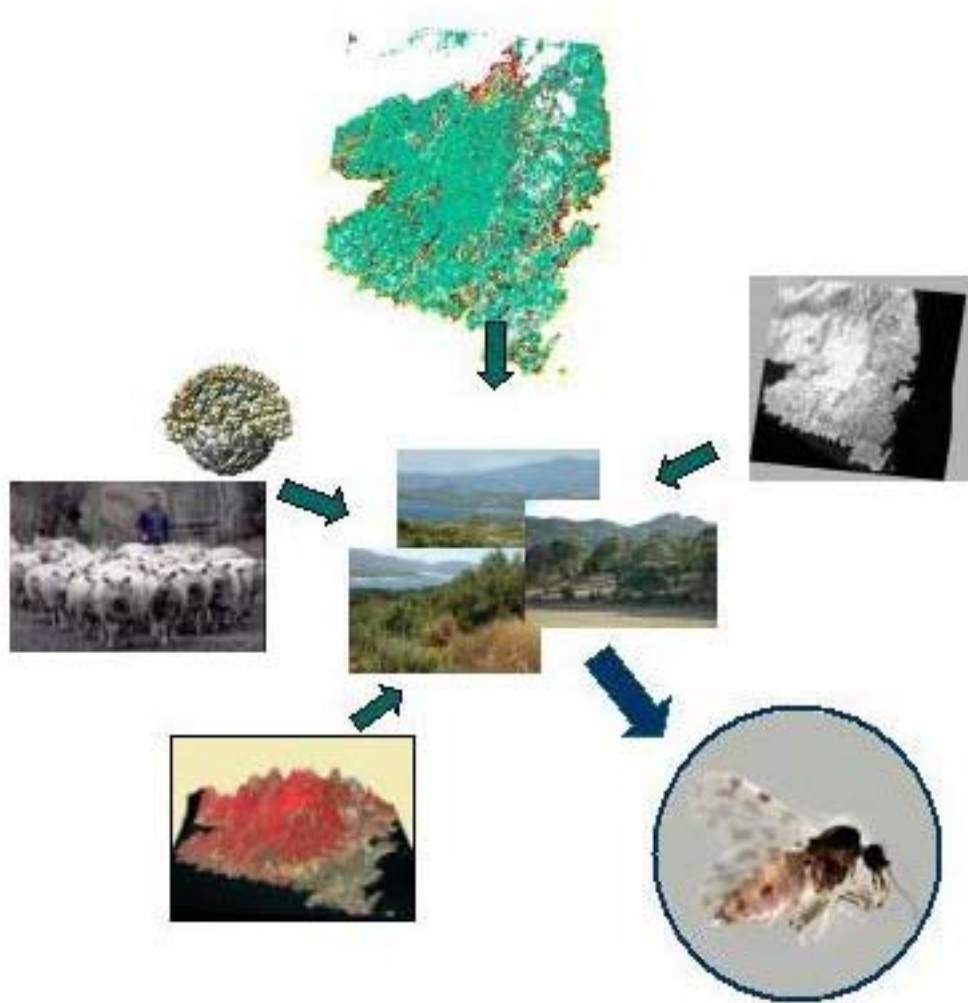
il faut qu'il ne fasse ni trop chaud, ni trop froid, qu'il pleuve assez, mais pas trop, qu'il y ait des endroits douilletts où se reposer et des moutons pas trop loin pour se nourrir.





Et voilà comment les détectives satellites retrouvent les petits mouchérons imicola. D'autres détectives satellites enquêtent sur d'autres criminels transmettant des maladies : les moustiques et le paludisme, les mouches tsé-tsé et la maladie du sommeil...
Merci à toi de m'avoir suivi dans cette investigation !

Image représentative du projet :



Résumé :

Du fait des changements climatiques et des actions de l'homme, de nombreux écosystèmes acquièrent la capacité d'héberger de nouveaux pathogènes majeurs. Ainsi, depuis quelques années, la répartition des vecteurs (particulièrement sensibles à leur écosystème) et des maladies qu'ils transmettent se modifie. L'objectif de ce projet, commandité par le Cirad-emvt, est d'étudier les conditions d'installation d'un vecteur dans un nouvel écosystème en prenant l'exemple de *Culicoides imicola*, le moucheron vecteur de la fièvre catarrhale ovine en Corse. Nous proposons ici d'utiliser la télédétection pour caractériser l'environnement favorable à *C. imicola*.

Pour réaliser ce projet, les données concernant les caractéristiques des élevages et leur statut vis-à-vis de la fièvre catarrhale ovine ont tout d'abord été organisées. Parallèlement, les paramètres caractérisant l'environnement et susceptibles d'influer sur la répartition du vecteur ont été obtenus à partir d'une image SPOT et d'un MNT. L'image satellitale a permis de calculer le NDVI et de classer la végétation selon deux méthodes (pixel à pixel et orientée objet). Du MNT ont été extraits la pente, l'altitude, l'ensoleillement et l'orientation. Fondés sur la comparaison du voisinage des élevages sains et infectés, des modèles logistiques expliquant la survenue de foyers de fièvre catarrhale ovine ont été construits. Ils font ressortir le rôle de la latitude, de certaines classes de végétation et du NDVI. Cependant, en raison de la qualité des données épidémiologiques initiales, ces modèles doivent être interprétés avec prudence.

Ce projet a néanmoins été l'occasion de mettre en oeuvre la démarche proposée initialement : construire un modèle expliquant la répartition des foyers de fièvre catarrhale ovine en Corse. Cette démarche, présentée aux partenaires impliqués dans le projet, devrait permettre d'ouvrir de nouvelles collaborations indispensables à la construction d'un modèle plus solide. Enfin, l'un des apports importants de ce projet est l'identification de nouvelles perspectives de recherche dans différents domaines (traitement d'image, analyse statistique...).

Abstract :

Due to climatic changes and human activities, many ecosystems are acquiring the capacity to shelter major pathogens. Because of the sensitivity of vectors to their environment, vector and vector-borne disease distributions have evolved greatly during the past years. The aim of this project, proposed by Cirad-emvt, is to study required conditions for the establishment of a vector in a new ecosystem, taking the example of *Culicoides imicola*, bluetongue vector, in Corsica. We propose to characterize "landscapes" favorable to *C. imicola* using remote sensing techniques.

To achieve this project, data concerning sheep farm characteristics and sanitary status were collected. In parallel, environmental factors likely to influence vector distribution were extracted from a SPOT image and from a DEM. From this image, NDVI was calculated and two methods of classification were carried out. The DEM enabled us to extract slope, altitude, sunshine, and orientation. We then compared the neighborhood of infected and disease-free farms to build a logistic model explaining bluetongue occurrence. These models highlighted the role of latitude, particular vegetation classes and NDVI. Nevertheless, due to poor quality of original epidemiological data, these models must be considered with caution.

Yet, this project enabled us to unfold the entire process initially proposed leading to the construction of a model explaining distribution of bluetongue in Corsica. This process, once presented to project partners could facilitate the new collaborations essential to put up a robust model. Finally, this project brought up new research prospects in many fields (image processing, statistic analysis...).

Mots clefs : télédétection, SIG, épidémiologie, maladie vectorielle, analyse spatiale, *Culicoides imicola*, fièvre catarrhale ovine, Corse

Key words : remote sensing, GIS, epidemiology, vector-borne disease, spatial analysis, *Culicoides imicola*, bluetongue, Corsica