

*Identification par télédétection  
des paysages méditerranéens  
favorables à *Culicoides imicola*,  
vecteur de la fièvre catarrhale  
ovine :  
validation de modèles statistiques  
de risque en Corse du Sud et dans  
le Var.*

**Bruno Barragué**

**Sous la direction d'Annelise Tran : Chercheur au CIRAD**

département Elevage et Médecine Vétérinaire, Unité Epidémiologie et  
Ecologie des Maladies animales

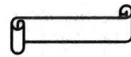
*Septembre 2006*

## *Remerciements*

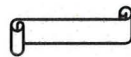
Mes premiers remerciements s'adressent à Annelise Tran pour m'avoir encadré avec autant de bienveillance et de sympathie.

Je remercie également Hélène Guis pour les mêmes raisons et pour la rigueur qu'elle a su, tant bien que mal, me transmettre dans mes méthodes de travail.

Merci à toutes les deux pour tout ce temps pris pour me conseiller et m'aider.

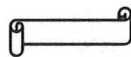


Je remercie aussi l'ensemble des membres de l'UR 16 du CIRAD pour leur accueil et leur collaboration.



Je remercie vivement l'ensemble des membres de la Maison de la Télédétection pour l'atmosphère chaleureuse et conviviale de travail au quotidien, dans un environnement exceptionnel.

Merci particulièrement à tous les stagiaires pour la bonne ambiance générale et les bons moments passés ensemble.





## Sommaire

<b>Introduction.....</b>	<b>4</b>
<b>1. Contexte général et objectifs du stage</b>	<b>6</b>
1.1. Organisme et structure d'accueil : le CIRAD et la Maison de la Télédétection.....	6
1.2. Etat des connaissances sur la maladie et le vecteur .....	6
1.2.1. Généralités.....	6
1.2.2. Présentation du travail d'Hélène Guis .....	7
1.3. Objectif du stage : validation des modèles statistiques établis précédemment.....	8
1.3.1. Présentation des modèles de risque de survenue de la FCO. ....	9
1.3.1.1. Modèles établis à 500 mètres. ....	9
1.3.1.2. Modèle établi à 1 km. ....	10
1.3.1.3. Modèle établi à 2 km. ....	10
1.3.2. Cas de la Corse : validation externe des modèles de risque.....	11
1.3.3. Cas du Var : essai de validation en zone indemne via des données entomologiques.....	12
<b>2. Approche opérationnelle</b>	<b>12</b>
2.1. Démarche.....	12
2.2. Les données relatives à la bluetongue.....	13
2.2.1. Les données épidémiologiques .....	13
2.2.2. Les données entomologiques .....	13
2.3. Méthode de cartographie de l'occupation du sol.....	14
2.3.1. Protocole.....	14
2.3.2. Caractéristiques des images satellitales .....	14
2.3.3. Acquisition des relevés de végétation. ....	15
2.3.4. Classification orientée objet : étapes et démarche. ....	16
2.3.4.1. La segmentation de l'image : définition des objets. ....	16
❖ Segmentation de l'image d' Ajaccio :.....	18
❖ Segmentation de l'image du Var.....	18
2.3.4.2. Principe de la classification supervisée orientée-objet par plus proche voisin. ....	19
2.3.5. Résultats des cartes obtenues.....	20
2.3.5.1. Classification de l'image SPOT - secteur Ajaccio .....	21
2.3.5.2. Classification de l'image SPOT – secteur Var .....	24
2.4. Extraction des variables environnementales .....	26
2.4.1. Extraction issue du MNT : l'ensoleillement moyen .....	26
2.4.2. Extraction des pourcentages de classe d'occupation du sol. ....	27
2.4.3. Extraction des variables paysagères.....	27
2.4.3.1. Définition .....	27
2.4.3.2. Méthode d'extraction .....	28
<b>3. Validation externe des modèles établis en Corse du Sud</b>	<b>29</b>
3.1. Calcul des probabilités de risque.....	29
3.2. Détermination d'une valeur seuil et évaluation du pouvoir discriminant des modèles de risque : analyse de ROC .....	30
3.3. Résultats de l'application externe des modèles.....	32
3.3.1. Validation dans le secteur d' Ajaccio .....	32
3.3.1.1. Modèles établis à 500 mètres : .....	32
3.3.1.2. Modèle établi à 1km :.....	33
3.3.1.3. Modèle établi à 2 km :.....	33
3.3.2. Validation dans le secteur du Var .....	34
3.3.2.1. Modèle établi à 1km :.....	34

**4. Discussion et perspectives 34**

- 4.1. Discussion sur les données relatives à la bluetongue..... 34**
- 4.2. Discussion sur les paramètres environnementaux ..... 36**
- 4.3. Discussion sur les modèles statistiques de risque ..... 36**

**Conclusion.....38**

**Bibliographie.....39**

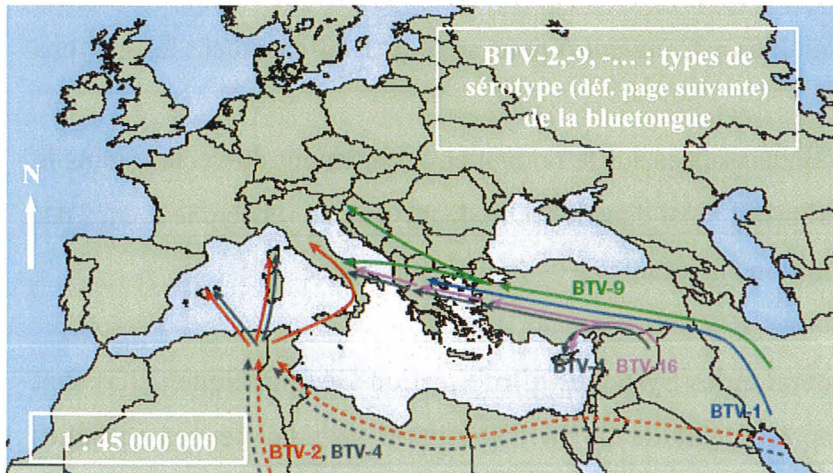
**Table des figures et tableaux.....41**

**Annexes.....42**



## Introduction

La répartition géographique des maladies dites émergentes<sup>1</sup> est en constante évolution du fait notamment du réchauffement climatique et de la modification des écosystèmes que cela induit. Ces maladies sont majoritairement à transmission vectorielle<sup>2</sup> du fait même de la sensibilité des vecteurs à leur environnement, dont tout changement peut modifier leur



**Figure 1 : progression de la FCO dans le bassin méditerranéen [Purse *et al.*, 2005]**

répartition et par conséquent celle des maladies qu'ils transmettent. C'est ainsi que la fièvre catarrhale ovine (FCO), maladie d'origine subsaharienne, est désormais présente dans le bassin méditerranéen.

La FCO (communément appelée bluetongue en anglais) est une maladie virale transmise par des moucheron hématophages du genre *culicoides* et affectant tous les ruminants sauvages et domestiques. L'infection ne provoque de signes cliniques que chez les moutons et a des conséquences économiques majeures pour l'élevage dans les régions touchées, tant en termes de santé animale que de répercussions sur les échanges commerciaux. La bluetongue est inscrite sur la liste A de l'Office International des Epizooties<sup>3</sup> (OIE<sup>4</sup>), impliquant d'importantes contraintes sur les ventes d'animaux (arrêt des échanges, etc.). Le moucheron *Culicoides imicola* est le vecteur principal de la bluetongue en Europe. Il s'agit d'un vecteur tropical, d'origine subsaharienne mais qui s'est installé récemment dans les pays du pourtour méditerranéen, notamment en France : en Corse depuis 2000 [Zientara *et al.*, 2000] et dans le Var très

<sup>1</sup> Une maladie émergente est une maladie infectieuse dont le taux d'incidence (nombre de nouveaux cas dans une population par unité de temps) a significativement augmenté dans les 20 dernières années, ou risque d'augmenter dans un avenir proche) [Tran *et al.*, 2005].

<sup>2</sup> Maladie vectorielle : qui est transmise par un arthropode hématophage qui assure la transmission biologique (ou mécanique) active d'un agent infectieux d'un vertébré à un autre vertébré [Rodhain *et al.*, 1985]

<sup>3</sup> Se dit d'une maladie qui frappe simultanément un grand nombre d'animaux (équivalent à « épidémique » pour une maladie humaine) [Garnier *et al.*, 1999]

<sup>4</sup> C'est l'organisation mondiale chargée de la santé animale.

récemment, depuis automne 2004. Depuis 1999, ce vecteur a généré plusieurs épizooties dans le bassin méditerranéen, notamment en Italie, Sardaigne, Corse, Tunisie, Algérie, Iles Baléares, Maroc, entraînant la perte de plus de 300 000 moutons. Suite à la confirmation de l'installation de populations de *Culicoides imicola* dans le Sud de la France (mai 2005), le risque pour la métropole est aujourd'hui majeur. Des vaccins existent mais sont lourds à mettre en place puisque 24 sérotypes<sup>5</sup> différents de FCO sont aujourd'hui connus [Baudoux *et al.*, VademecumFCO]. Les caractéristiques des écosystèmes dans lesquels évoluent aujourd'hui ce vecteur sont très différentes de celles dont il est originaire il devient nécessaire de cibler les zones à risque (favorables à *C. imicola*) pour appliquer efficacement les moyens de lutte.

A ce jour, il y a peu de connaissances sur le nouvel environnement du vecteur dans le bassin méditerranéen ; ce stage s'inscrit ainsi dans le cadre de travaux de recherche menés par le cirad pour caractériser les conditions favorables à *C. imicola* en France.

Les données de télédétection et les Systèmes d'Information Géographique (SIG) sont des outils particulièrement adaptés à l'étude de la distribution spatiale de ces maladies vectorielles. Les classifications d'images satellites SPOT 5, par exemple, permettent de réaliser des cartographies de l'occupation du sol à une échelle très fine (résolution de 10 mètres) et donc l'identification de paramètres environnementaux favorables au vecteur (via la spatialisation de données relatives à celui-ci dans un SIG).

La première partie de ce rapport expose le contexte général de l'étude ainsi que les objectifs. L'approche opérationnelle pour mener cette étude est ensuite détaillée (partie 2) avant la phase de validation des modèles<sup>6</sup> de risque de survenue de la maladie établi en Corse du sud (partie 3). Enfin, la dernière partie est consacrée à la discussion sur cette étude et aux perspectives d'amélioration.

---

<sup>5</sup> Sérotype : catégorie dans laquelle on classe les virus et les bactéries selon leurs réactions en présence de sérums contenant des anticorps spécifiques [Garnier *et al.*, 1999].

<sup>6</sup> Modèles statistiques établis par Hélène Guis (doctorante en épidémiologie au cirad).



# 1. Contexte général et objectifs du stage

## 1.1. **Organisme et structure d'accueil : le CIRAD et la Maison de la Télédétection**

Cette étude est commanditée par le département Elevage et Médecine Vétérinaire (EMVT) du Centre de coopération International en Recherche Agronomique pour le Développement (CIRAD). Plus exactement, ce stage prend place au sein de l'Unité de Recherche (UR) 16, Epidémiologie et Ecologie des Maladies Animales, regroupant aussi des compétences en entomologie et en géomatique pour la surveillance de maladie comme la fièvre de la vallée du rift, la fièvre du Nil occidental, la grippe aviaire. Cette unité est donc en charge de la gestion de la crise de la FCO auprès de la Direction Générale de l'Alimentation (DGAL), au Ministère de l'agriculture.

Plusieurs équipes sont ainsi en étroite collaboration et ce travail prend place au sein de l'Unité Mixte de Recherche (UMR) Territoires, Environnement, Télédétection et Information Spatiale (TETIS) pour les aspects de traitement d'image et de SIG. Cette équipe est localisée à la Maison de la Télédétection, structure regroupant des équipes de divers centres de recherches (Cirad, Cemagref<sup>7</sup>, Engref<sup>8</sup>, IRD<sup>9</sup>) qui travaillent avec les outils géomatiques.

## 1.2. **Etat des connaissances sur la maladie et le vecteur**

### 1.2.1. Généralités

Les symptômes affectant les ovins se manifestent par une fièvre, une inflammation de la muqueuse buccale, des oedèmes, des raideurs musculaires et des boiteries (annexe 1). Elle peut entraîner la mort de plus de 20 % du cheptel dans les zones infectées. L'infection est généralement asymptomatique chez les bovins et les caprins mais ces espèces pourraient jouer le rôle de réservoirs<sup>10</sup> dans la transmission de la maladie. Aujourd'hui, 24 sérotypes de virus sont connus, mais ils ne présentent pas de protection croisée. Un vaccin ne protège donc que

---

<sup>7</sup> Institut de recherche pour l'ingénierie de l'agriculture et de l'environnement

<sup>8</sup> Ecole Nationale du Génie Rural des Eaux et Forêt

<sup>9</sup> Institut de Recherche et de Développement.

<sup>10</sup> Réservoir : « 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 » [Rodhain *et al*, 1985]



pour le sérotype contre lequel il est dirigé, néanmoins, il existe des vaccins regroupant plusieurs sérotypes [Guis, 2004].

*Culicoides imicola* (annexe 2), vecteur de la FCO, est d'origine sub-saharienne et se trouve aujourd'hui dans nos latitudes grâce à une dispersion progressive et passive par les vents d'altitudes sur plusieurs centaines de kilomètres. A l'inverse, il a une dispersion active faible, estimée à quelques centaines de mètres autour de son lieu de reproduction. Peu d'éléments sur sa bio-écologie sont connues : les larves préféreraient des sols riches en matières organiques avec un niveau d'humidité suffisant. Les femelles adultes, crépusculaires, effectuent plusieurs repas de sang au cours de leur vie, et peuvent ainsi transmettre le virus à plusieurs 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) [Guis, 2004].

### 1.2.2. Présentation du travail d'Hélène Guis

Le travail de thèse d'Hélène Guis consiste à caractériser l'environnement favorable à *C. imicola* en Corse. Elle a donc établis divers modèles statistiques (présentés dans la partie suivante) développés à partir d'image SPOT afin d'expliquer la répartition des foyers de FCO et dont peuvent être issus des cartes d'aléa. L'originalité de son approche repose sur un travail à partir d'image satellite à haute résolution spatiale (10 mètres) permettant le développement de modèles à échelle fine et l'intégration de la notion de paysage comme variable dans les modèles. Les modèles déjà existant sur la répartition et l'abondance des culicoides s'appuient sur des variables météorologiques [Ward, 1996] [Wittmann *et al*, 2001] et plus récemment sur des indices extraits d'image satellite comme le NDVI<sup>11</sup> [Tatem *et al*, 2003]. Elle a ainsi mis en évidence des facteurs de risque tels que la latitude et la mixité<sup>12</sup> des élevages, et des facteurs protecteurs tels que, le NDVI, l'ensoleillement moyen, les cultures, le maquis, la garrigue et la forêt. De plus, des variables paysagères (relatives à la structure du paysage) ont été intégrées dans la plupart de ces modèles.

---

<sup>11</sup> Normalised Difference Vegetation Index. Définition § 2.3.5.

<sup>12</sup> Est entendu par mixité la présence d'au moins deux espèces dans l'exploitation agricole (bovins, ovins, caprins)

Ces variables ont été extraites autour des élevages recensés et localisés par GPS (Global Positioning System) à trois échelles différentes : 500 m, 1 et 2 km. Le choix de ces échelles repose sur deux considérations :

- la dispersion active de *C. imicola* est en moyenne de l'ordre de quelques centaines de mètres.
- les élevages ayant été pointés au niveau de la bergerie, il a été estimé que les pâtures se trouvent dans un rayon de quelques centaines de mètres.

Ainsi, suivant l'échelle adoptée, les modèles intègrent différentes variables.

La recherche des variables potentiellement explicatives du statut sanitaire des élevages s'est faite par une régression logistique, modèle permettant d'exprimer la relation entre une variable qualitative à deux classes (élevage sain ou infecté) et des variables pouvant être qualitatives ou quantitatives. L'équation suivante décrit le modèle logistique [Guis, 2004] :

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

*P* : probabilité d'être infecté  
*α* : constante  
*β* : coefficient (constant)  
*X* : variable explicative

### **1.3. Objectif du stage : validation des modèles statistiques établis précédemment**

Ce travail s'inscrit dans la thèse réalisée par H. Guis. Il a pour objectifs d'évaluer ces modèles établis en Corse du Sud sur la probabilité de survenue d'un foyer de FCO, et de poursuivre ainsi l'étude de la caractérisation de l'environnement favorable aux *Culicoides*, vecteur de la maladie, dans le bassin méditerranéen. Ces modèles ont été établis via la description du milieu au voisinage des élevages, à partir de l'hypothèse que les paramètres environnementaux (topographiques, type de végétation, organisation du paysage, ...) peuvent être identifiés à partir d'images satellitaires. Les modèles existants de répartition du vecteur ont tous été élaborés à partir d'images à basse résolution spatiale. L'originalité des modèles à valider pour cette étude repose sur l'intégration de la notion de paysage possible grâce au travail sur des images satellites à haute résolution spatiale (10 mètres). Une échelle aussi fine permet de mieux appréhender le nouvel écosystème du vecteur.

La validation porte sur le secteur d'Ajaccio (juste au nord du secteur dont sont issus les modèles de répartition des foyers – fig2, p11) et sur le Var où le vecteur a donc récemment fait son apparition.

De plus, ces modèles ont été élaborés avec les données sur la répartition des foyers, en partant de l'hypothèse que les statuts infectés ou indemnes des élevages sont essentiellement conditionnés par un environnement favorable ou non à *C. imicola*.

### 1.3.1. Présentation des modèles de risque de survenue de la FCO.

Quatre modèles statistiques ont donc été établis par Hélène Guis via l'extraction et l'analyse de variables environnementales autour des élevages (indemnes et infectés) recensés, grâce à un long travail de recueil de l'information sur l'état sanitaire des élevages, dans la partie sud de la Corse. Trois de ces quatre modèles intègrent des variables paysagères. L'extraction des variables s'est faite dans des zones tampons, autour des élevages, à trois échelles : 500 mètres (deux modèles), 1 km et 2 km. Ces modèles sont rapidement présentés dans cette partie mais seront plus détaillés (notamment la description des variables) dans la partie 2.4 consacrée à l'extraction des paramètres environnementaux incluses dans les modèles. Ils ont été élaborés via une régression logistique<sup>13</sup> pas à pas ascendante afin de ne retenir que les variables significatives et explicatives du statut sanitaire des élevages. Ces modèles logistiques se présentent par l'équation suivante :

#### 1.3.1.1. Modèles établis à 500 mètres.

Deux modèles ont donc été réalisés à cette échelle. Ils se différencient par la prise en compte de variables paysagères par l'un d'eux.

- Modèle sans variable paysagère :

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

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

*C* : constante

*MIXITE* : élevage où l'on trouve au moins deux espèces d'animaux

*Y* : latitude

*Culture & MAQUIS/GARRIGUE/FORET* : classes de végétation

<sup>13</sup> Ce modèle permet d'exprimer la relation entre une variable qualitative à deux classes (ici élevage sain ou infecté) et des variables pouvant être qualitatives ou quantitatives.



- Modèle avec variables paysagères :

$$\ln(P/(1-P)) = C + a \text{ MIXITE} + b Y + c \text{ Culture} + \text{LS PRD}$$

$$\text{Avec } C = 513,05878 \quad a = 4.057409867 \quad b = -0,000124649$$

*LS PRD* (variable paysagère) : Landscape Patch Richness Density (les variables paysagères seront détaillées dans la troisième partie)

### 1.3.1.2. Modèle établi à 1 km.

$$\ln(P/(1-P)) = C + a \text{ MIXITE} + b \text{ PELOUSE} + c \text{ ZNV\_LSI} + \text{Prairie\_LSI} + \text{MGF\_LSI} + d \text{ Ensoleillement moyen}$$

$$\text{Avec } C = -3.367636666 \quad a = 4,892553996 \quad b = 1,647982495 \quad c = 0,422961652 \\ d = -9,024275285$$

*ZNV\_LSI* : Zones Non Végétalisées\_Landscape Shape Index  
*MGF\_LSI* : Maquis Garrigue Forêt\_Landscape Shape Index

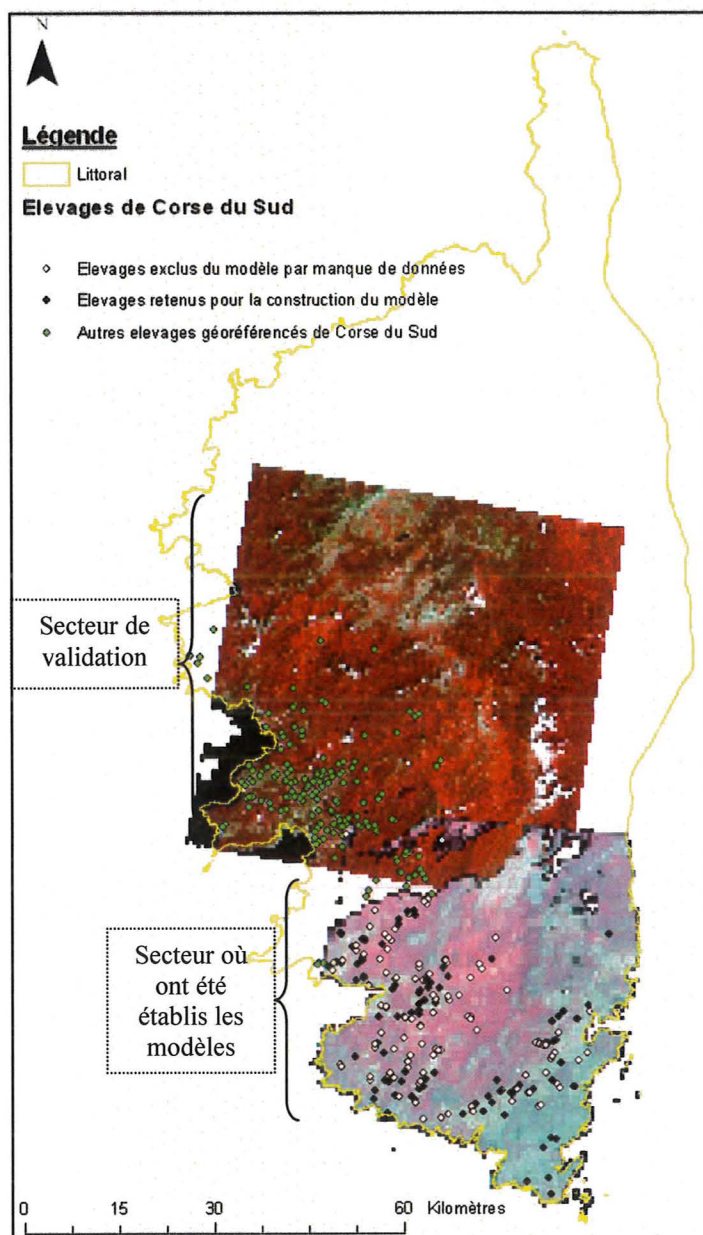
### 1.3.1.3. Modèle établi à 2 km.

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

$$\text{Avec } C = 612,863 \quad a = 3,615126990 \quad b = -0,000147644 \quad c = 0,215908177 \quad d = -13,3067$$

*Prairie\_NP* : Prairie\_Nombre de Patch

### 1.3.2. Cas de la Corse : validation externe des modèles de risque



Le premier volet de cette étude consiste à tester les différents modèles établis en Corse du sud sur le secteur d'Ajaccio qui se juxtapose à ce premier. La maladie est également présente dans ce secteur, les modèles pourront ainsi être appliqués à partir de données épidémiologiques similaires à l'établissement de ceux-ci.

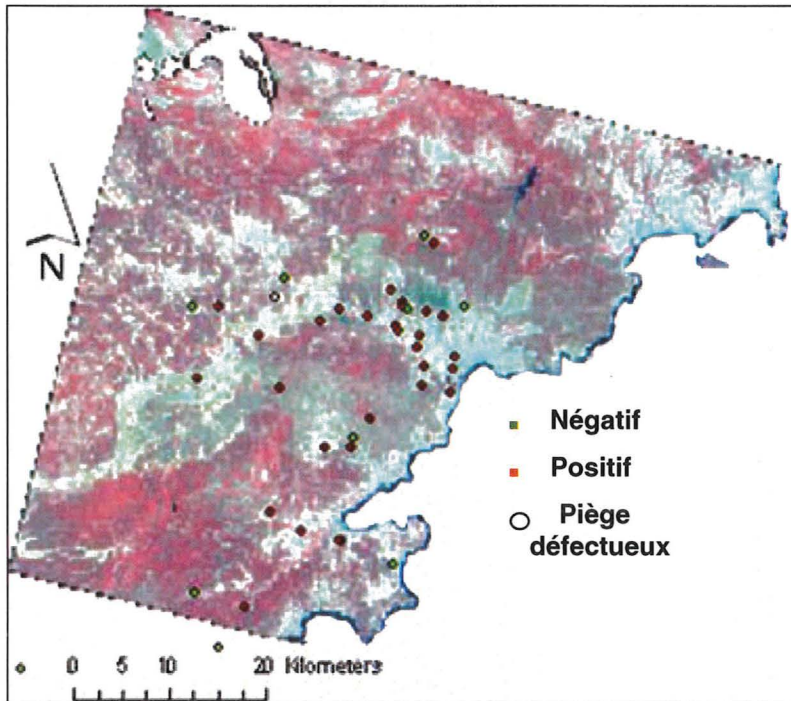
**Figure 2 : Localisation des élevages de Corse du sud sur les images satellitaires SPOT**

*Données SPOT / Programme ISIS<sup>14</sup>, © CNES<sup>15</sup> (2001). [Guis, 2004]*

<sup>14</sup> ISIS : Incitation à l'utilisation Scientifique d'Images SPOT

<sup>15</sup> CNES : Centre National d'Etude Spatiale

### 1.3.3. Cas du Var : essai de validation en zone indemne via des données entomologiques



Le contexte est différent dans le Var. En effet, le vecteur est bien présent (depuis 2004) mais il n'y a pas encore d'élevage infecté recensé à ce jour. L'essai de validation des modèles se fera donc par rapport aux relevés entomologiques, dont l'hypothèse est faite qu'un piège positif correspond à un élevage infecté et inversement.

Figure 3 : Localisation des relevés entomologiques dans le Var sur l'image satellite

## 2. Approche opérationnelle

### 2.1. Démarche.

La validation des modèles et le calcul des probabilités de risque d'infection nécessitent l'extraction des variables qui les composent. Le schéma suivant présente les étapes nécessaires à cette opération.

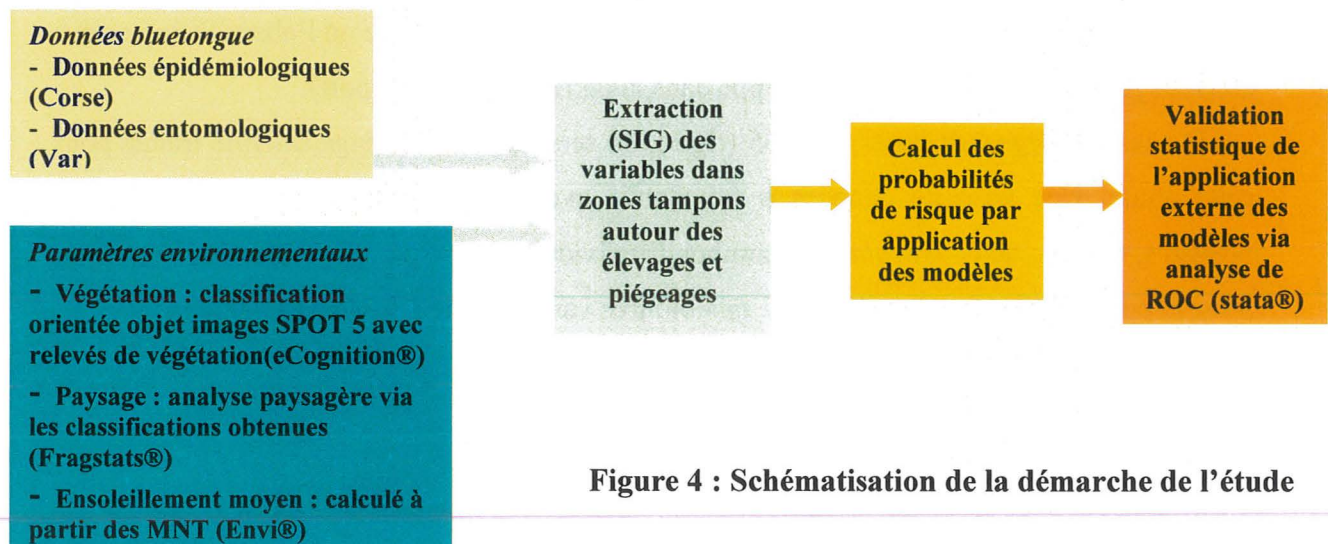


Figure 4 : Schématisation de la démarche de l'étude



## **2.2. Les données relatives à la bluetongue**

### **2.2.1. Les données épidémiologiques**

Ces données sont relatives au secteur de la Corse, la maladie n'étant pas présente dans le Var. L'ensemble des informations sur les élevages a été recueilli par Hélène Guis, puis organisé et renseigné sur leurs caractéristiques (espèces élevés, mixité au sein des élevages, statut sanitaire des animaux). Ce travail a demandé un temps long du fait de l'hétérogénéité des données disponibles. Seule l'information sur la taille des cheptels n'a pu être recueillie de façon assez exhaustive pour être utilisée.

Ces informations sont intégrées dans une base de données géoréférencée. Ainsi, dans le secteur étudié pour ce stage ce sont 151 élevages qui sont recensés, renseignés et cartographiés (fig.2). Ils serviront pour la validation des modèles de répartition des foyers de fièvre catarrhale ovine établis dans la partie sud de l'île.

La mixité des élevages est une variable commune à l'ensemble des modèles. Le rôle de réservoir joué par les espèces bovines et caprines dans la diffusion de la maladie étant, à priori, déterminant.

### **2.2.2. Les données entomologiques**

Ces données sont utilisées pour l'essai de validation dans le Var, puisque ce sont les seules données sur lesquels il est possible de tester les modèles de risque, la maladie n'étant pas encore survenue en métropole. Il est important d'insister sur la notion d' « essai » puisque plusieurs hypothèses doivent être préalablement posées pour justifier de transposer des modèles fondés sur la répartition de foyers, sur désormais, la présence ou l'absence du vecteur dans les pièges<sup>16</sup> (cet aspect est développé dans la partie 3.1 de ce rapport).

Sur les 42 piégeages posés en 2005 (fig.3), 38 se situent sur l'image satellite traitée pour cette étude. Une base de données a été établie avec les coordonnées géographiques, la présence, ou non, de *C. imicola* et le nombre d'individus capturés par piège. Sur ces 38 pièges, 30 sont positifs avec des densités de *C. imicola* très variables (de 1 à 3186 individus).

---

<sup>16</sup> Un exemple de piège utilisé pour la capture de *C. imicola* est présenté en annexe 3.

## 2.3. Méthode de cartographie de l'occupation du sol

### 2.3.1. Protocole

L'organisation schématique de la réalisation de ces cartes de végétation est présentée dans la figure suivante :

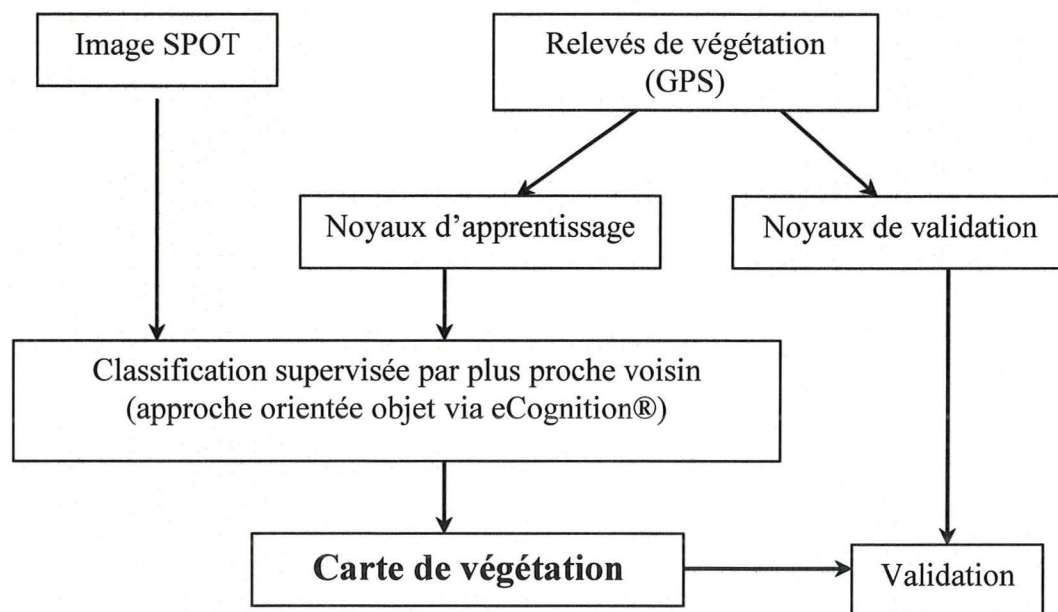


Figure 5 : Schéma de la méthode de cartographie de l'occupation du sol

### 2.3.2. Caractéristiques des images satellitaires

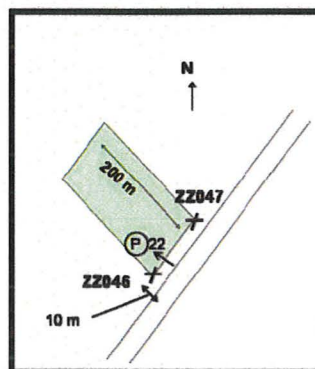
Deux images SPOT 5, une pour le secteur d'Ajaccio et l'autre pour la partie littorale du Var ont été acquises pour la réalisation des cartes de végétation (*programme ISIS, CNES*). Ces images ont une résolution de 10 mètres avec 4 bandes spectrales : PIR (Proche Infrarouge), rouge, vert et MIR (Moyen Infrarouge). Leurs caractéristiques, sensiblement différentes, sont présentées le tableau 1.

<i>Caractéristiques</i>	<i>Image Ajaccio</i>	<i>Image Var</i>
Numéro de l'image	5059265020720124282J	5 053-262/5 05/07/11 10 :18 :16 1 J
Satellite (capteur)	SPOT 5 (HRG2)	SPOT 5 (HRG1)
Produit	Spotview (niveau 3 ortho : l'image est corrigé géométriquement et projetée)	Spot Scène (niveau 2A : l'image est corrigé géométriquement et projeté)
Date de prise de vue	20/07/2002	11/07/2005
Bandes spectrales (µm)	0.78 - 0.89 (PIR), 0.61 - 0.68 (rouge), 0.5 - 0.5 (vert), 1.58 - 1.75 (MIR)	0.78 - 0.89 (PIR), 0.61 - 0.68 (rouge), 0.5 - 0.5 (vert), 1.58 - 1.75 (MIR)
Taille du pixel	10 mètres	10 mètres
Taille de l'image	6953 lignes x 6953 colonnes	7498 lignes x 8338 colonnes
Projection (datum, ellipsoïde de référence)	Lambert IV carto Paris (NTF <sup>17</sup> , ellipsoïde de clarke 1880 IGN)	UTM WGS84 par défaut (NTF, ellipsoïde de clarke 1880 IGN)

**Tableau 1 : Caractéristiques des images SPOT 5**

### 2.3.3. Acquisition des relevés de végétation.

Les données sur la végétation de Corse sont issues de 132 relevés terrain acquis lors d'une mission réalisée en septembre 2005. Pour chacun d'eux est décrite l'occupation du sol via une fiches de relevés de végétation où sont classées les groupes végétaux de Corse, leurs espèces respectives, leur localisation (par au moins deux points GPS – ZZ046 et ZZ047 sur la fig.6) et un schéma permettant le positionnement du relevé effectué. Cette description est accompagnée au minimum d'une photo de la parcelle avec l'orientation de la prise de vue (P22). Les zones choisies répondaient au compromis d'être à la fois les plus grandes et les plus homogènes possibles.



**Figure 6 : Exemple de schéma permettant de positionner les relevés de végétation**

<sup>17</sup> NTF : Nouvelle Triangulation Française



Les photos ont souvent permis de mieux appréhender la réalité terrain lors de la définition des classes d'apprentissage et de validation pour la classification de l'image, étant donné que ces relevés ont été réalisés avant ce stage.

Les relevés de végétation dans le Var (au nombre de 93) ont été acquis au cours de ce stage, lors d'une mission de trois jours dans ce département. Le protocole suivi fut sensiblement le même que pour la Corse (fiche de relevés de végétation en annexe 1). Des secteurs particuliers ont été présélectionnés pour effectuer ces relevés :

- Tout d'abord, les secteurs proches des pièges entomologiques posés en 2005 furent privilégiés.
- Des zones homogènes d'un point de vue radiométrique ont été identifiées : une ACP<sup>18</sup> de l'image SPOT a été réalisée de façon à obtenir une image plus contrastée, suivie d'une classification non supervisée permettant de bien apprécier les zones homogènes radiométriquement. Un document avec des cartes papiers a ensuite été produit reprenant cette dernière classification ainsi qu'une composition colorée en fausses couleurs, découpées suivant les cartes IGN 1/25000 du secteur, et cartographié sur une grille géoréférencée. Ceci a facilité, sur le terrain, la localisation de zones homogènes dont l'itinéraire pour y accéder était ensuite repéré sur carte IGN.

Une base de données sous Excel a ensuite été créée avec les relevés pour la Corse et ceux acquis dans le Var (en format « dbf ») de façon à pouvoir les intégrer dans un SIG et les cartographier sur les deux images.

### **2.3.4. Classification orientée objet : étapes et démarche.**

Le logiciel eCognition® est ici utilisé pour réaliser ce type de classification. Le principe consiste, non pas à classer des pixels mais des objets homogènes. Ainsi, une première étape de segmentation précède la classification de l'image.

#### *2.3.4.1. La segmentation de l'image : définition des objets.*

La segmentation consiste à créer une image non classifiée, subdivisée en polygones élémentaires, les objets. Ceux-ci sont issus de l'agrégation des pixels pour lesquels

---

<sup>18</sup> ACP : Analyse en Composantes Principales : technique mathématique permettant de réduire un système complexe de variables corrélées entre elles en un plus petit nombre de dimensions..

l'augmentation de l'hétérogénéité (radiométrie et forme) est la plus faible, et ce jusqu'à qu'un seuil d'hétérogénéité, fixé par l'utilisateur soit dépassé.

Ainsi, deux principaux critères sont à renseigner pour effectuer la segmentation :

- ❖ *Le critère d'échelle* : il paramètre la taille des polygones. Ainsi, plus le paramètre d'échelle est élevé, plus les objets produits sont importants.
- ❖ *Le critère d'homogénéité* : celui-ci dépend de deux séries de paramétrage.

Tout d'abord, l'homogénéité résulte d'une pondération entre un critère spectral et un critère de forme. Le critère de couleur est dans la plupart des cas le plus important puisqu'il définit le seuil d'homogénéité des valeurs spectrales pour créer les objets.

Ensuite eCognition® permet d'affiner l'homogénéité des objets selon leur forme par une pondération entre deux critères, « le lissage » et la « compacité ». Cette seconde pondération permet d'affiner la segmentation de l'image en tenant compte de la forme lissée ou compacte des objets.

Ces deux paramétrages définissent la priorité qui est donnée à chacun des critères pour obtenir la segmentation la plus appropriée.

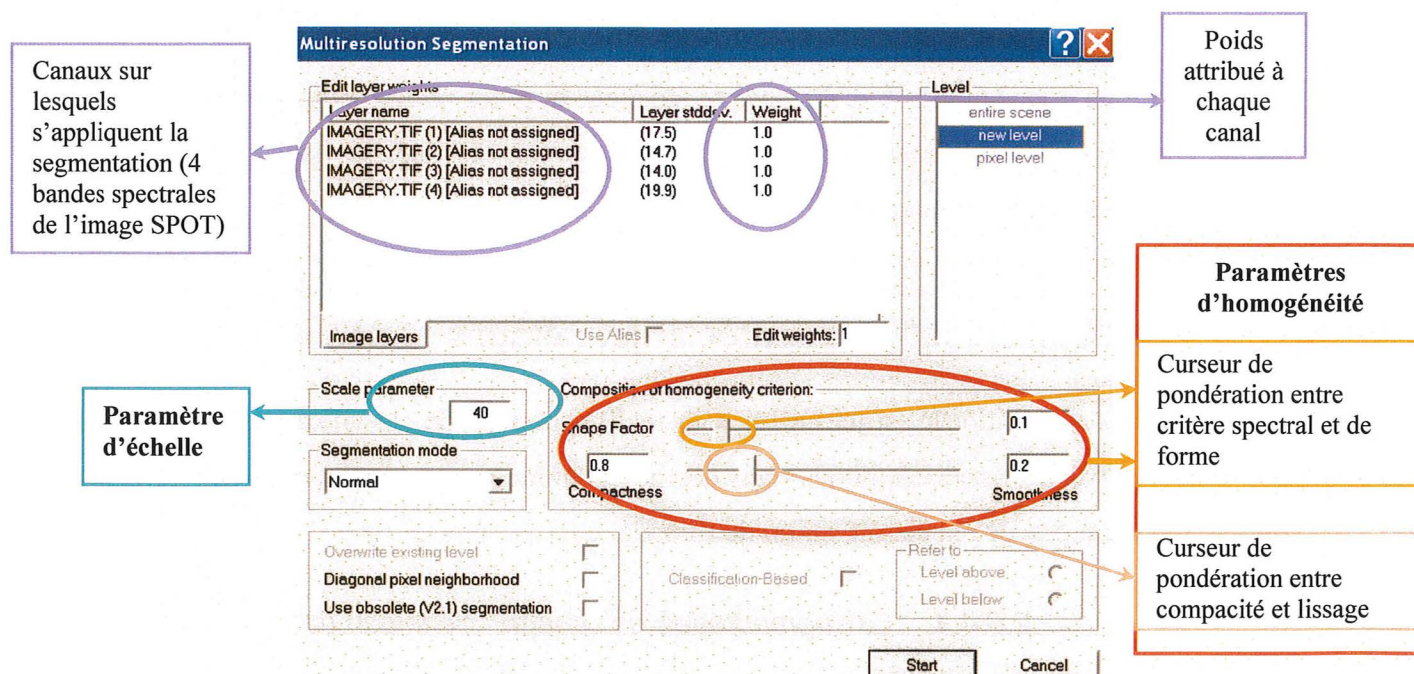


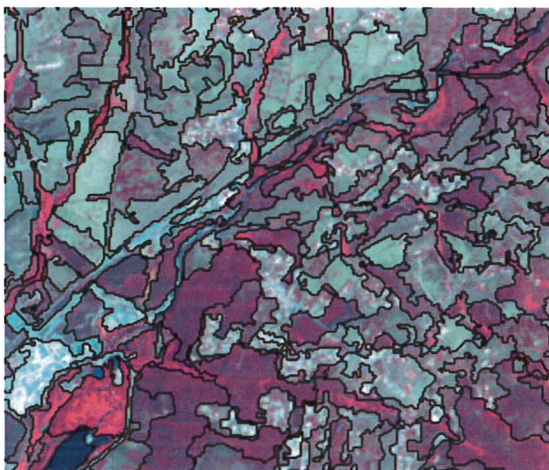
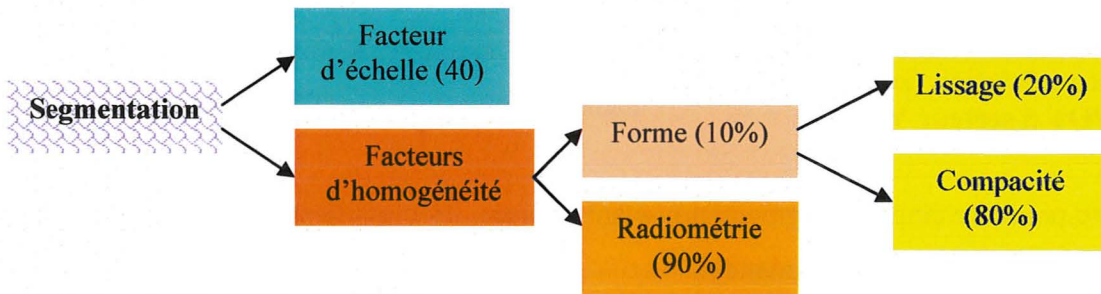
Figure 7 : Interface de paramétrage de la segmentation de l'image dans eCognition®



Pour obtenir une bonne segmentation de l'image, il est recommandé de lancer différentes segmentations avec différents paramètres jusqu'à ce que le résultat soit satisfaisant. Il est ainsi possible de travailler dans un premier temps sur un échantillon de l'image pour la définition des paramètres de manière à accélérer les temps de traitements avant de l'appliquer à l'ensemble de l'image.

❖ *Segmentation de l'image d'Ajaccio :*

Après plusieurs essais, les paramètres suivant ont été choisis :



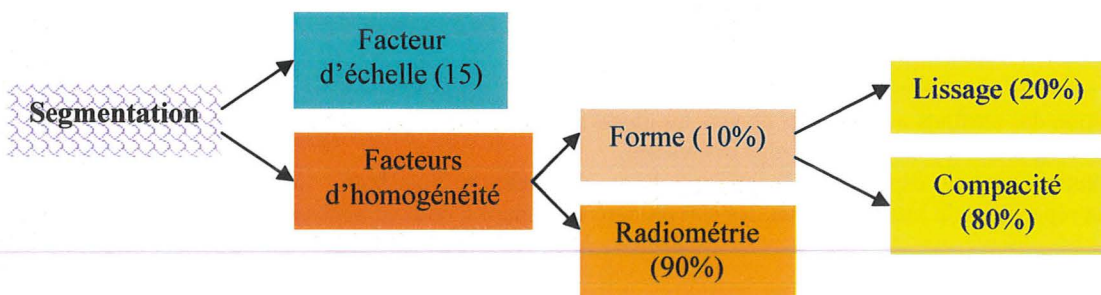
700 m

Figure 8 : Illustration des contours des objets après segmentation (zoom de l'image SPOT 5 secteur Ajaccio)

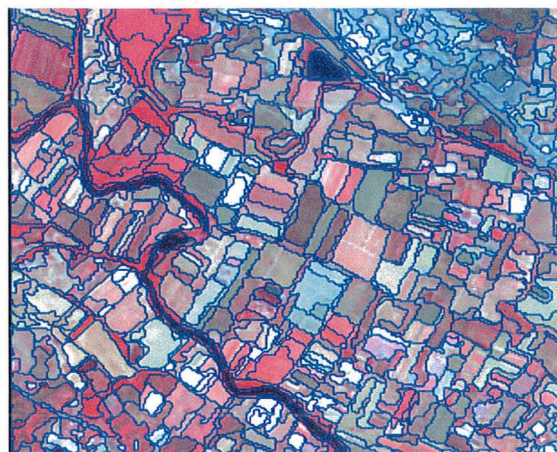
Le niveau de détail est suffisamment élevé, malgré un facteur d'échelle relativement important. L'occupation du sol sur cette partie de la Corse étant assez homogène avec peu d'effet de « mosaïque », justifie ce paramétrage. Le facteur spectral a été fixé à 90% car les objets à caractériser se distinguent plus par leur radiométrie que par leur forme. Pour les facteurs de forme, lissage et compacité, le choix s'est fait par évaluation visuelle.

❖ *Segmentation de l'image du Var*

De la même façon, les paramètres ci-après ont été sélectionnés :







**Figure 9 : Illustration des contours des objets après segmentation (zoom de l'image SPOT 5 secteur Var)**

Le facteur d'échelle (15) choisi ici est beaucoup plus fin que pour le secteur précédent. En effet, l'activité agricole dans la vallée de l'Argens étant importante, l'occupation du sol se caractérise par une mosaïque de parcelles facilement identifiables visuellement sur l'image (composition colorée en fausse couleur PIR, rouge, vert) zoomée. Ce paramètre d'échelle est apparu le plus pertinent pour bien séparer les parcelles entre-elles. Les facteurs d'homogénéité ont été arrêtés selon les mêmes critères que pour l'image de Corse.

Cette première étape de segmentation étant réalisée, les classifications des images par discrimination des objets peut commencer.

#### 2.3.4.2. Principe de la classification supervisée orientée-objet par plus proche voisin.

Des parcelles d'apprentissage et de validation, issues pour cette étude des relevés de terrain, représentatives de chaque classe sont définies. Ensuite, ECognition® présente l'intérêt de pouvoir employer un grand nombre d'attributs (en plus des valeurs spectrales) liés à la forme, la texture ou le contexte, pour décrire les objets de l'image. Ces attributs sont affectés aux différentes classes afin de créer un espace multidimensionnel (autant de dimensions que d'attributs). Chaque classe issue des parcelles d'apprentissage est alors caractérisée spécifiquement dans cet espace. Ainsi, lors de la classification, chaque objet de l'image est assigné à la classe la plus proche de lui dans l'espace des attributs.

Cette partition de l'espace dimensionnel est donc définie par l'utilisateur qui choisit les attributs à utiliser pour effectuer la classification. Ce choix peut se faire de manière empirique en testant la pertinence des attributs (fonction *feature view*<sup>19</sup> et fonction *2D feature space plot*<sup>20</sup> - annexe 1 -), ou de façon automatisée grâce à la fonction *Feature space*

<sup>19</sup> « visualisation des attributs » : il permet de voir les objets affichés selon la valeur qu'ils ont dans cet attribut et donc de faciliter le choix des attributs les plus intéressants.

<sup>20</sup> « représentation dans un espace à 2 dimensions d'une combinaison d'attributs » : cette fonction permet de représenter la répartition du nuage de point des classes dans un espace à deux dimensions (un attribut en X et un en Y).

*optimisation*<sup>21</sup>. Cette dernière permet de garder les attributs qui séparent au mieux les différentes classes en calculant la plus forte séparation entre celles-ci parmi les attributs préalablement définis par l'utilisateur.

Les classifications sous eCognition® sont fondées sur une logique floue défini par un seuil d'appartenance, entre 0 et 1, exprimant l'appartenance d'un objet à une classe. L'algorithme de classification de ce logiciel est de type « doux » (*soft classifier*) ce qui permet d'exprimer des incertitudes dans la description des classes, ainsi que le degré d'appartenance d'un objet dans plusieurs classes. En résumé, un objet sera affilié à la classe la plus proche dans l'espace multidimensionnel en fonction de sa valeur d'appartenance si celle-ci dépasse le seuil défini par l'utilisateur (fonction *edit minimum membership value*<sup>22</sup>). Par défaut, le seuil est fixé à 0,1. Différents seuils ont été testés pour les classifications des deux images. Le nombre d'objet non classé augmente avec l'attribution d'un seuil se rapprochant de 1. Après observation, les objets non classés avec des seuils élevés se retrouvent dans des classes cohérentes dans des classifications avec des seuils faibles. Le choix de garder le seuil par défaut de 0,1 a donc été préféré.

### 2.3.5. Résultats des cartes obtenues

Les classifications ont été effectuées sur des images satellites prises durant la période estivale. Les relevés de terrain ne datent pas de la même année ni de la même période ce qui a amené à interpréter certains relevés (particulièrement pour les zones cultivées dans le Var) et à en écarter d'autres (réponse spectrale de certains objets aberrantes par rapport au relevé terrain probablement du à des changements d'occupation du sol). Pour rappel, l'image de Corse date du 20 juillet 2002 et les relevés terrain de début septembre 2005. L'image du Var date du 11 juillet 2005 et les relevés de mai 2006.

Avant la phase de classification, les nuages, leurs ombres et la mer ont été masqués<sup>23</sup>. De plus, un indice de végétation, le NDVI<sup>24</sup>, a été calculé pour faciliter la classification et la discrimination des différents types de végétation. Cet indice de végétation à différence normalisée 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 :

$$\text{NDVI} = (\text{PIR} - \text{Rouge}) / (\text{PIR} + \text{Rouge})$$

<sup>21</sup> « optimisation d'un espace d'attributs »

<sup>22</sup> « éditer la valeur minimale d'appartenance »

<sup>23</sup> Ces opérations ont été réalisées sous Envi©

<sup>24</sup> NDVI : Normalised Difference Vegetation Index. Il aide la discrimination des classes de végétation.



L'évaluation de la qualité de ce type de classification consiste ensuite en une confrontation avec des vérités terrain (parcelles d'évaluation) par une analyse quantitative via une matrice de confusion. Celle-ci présente les classes de validation en colonnes et la classification en lignes (figure 10). La diagonale correspond aux pixels bien classés. Dans les matrices suivantes, les valeurs ont été pondérées en pourcentage de façon à attribuer le même poids à la validation de chaque classe. Ainsi, les indices d'appréciation globale de qualité, qui calculent la proportion de pixel correctement classée, ne sont pas biaisés par la disproportion de pixels que compte les diverses classes de validation. Deux indices de qualité sont ainsi donnés avec la matrice de confusion effectuée avec eCognition® :

- « Overall Accuracy » : correspond à la proportion de l'ensemble des pixels bien classés (d'où l'intérêt de la pondération).
- Indice de Kappa (KIA) : il permet une évaluation globale en tenant compte des erreurs en lignes et en colonnes [Girard M.C *et al.*, 1999]. Les classes ayant été pondérées, cet indice ne sera pas présenté dans les matrices suivantes.

### 2.3.5.1. Classification de l'image SPOT - secteur Ajaccio

Une première classification a été réalisée avec un espace multidimensionnel défini de 14 attributs, essentiellement de texture, (sélectionnés parmi plus de 90) offrant la meilleure séparabilité entre les classes d'après la fonction *feature space optimisation* (cf annexe 6). La matrice de confusion, ci-après, révèle une qualité et une précision peu satisfaisante, la classification ne sera donc pas présentée.

<b>Vérité terrain</b> <b>Classification</b>	Pinède	Zones non végétalisées	Prairie	Prairie_arborée-arbres isolés	cultures	Pelouse	Maquis-garrigue-forêt_de_feuillus	Eau	Total
Pinède	100	0	0	0	0	0	4,2	0	10,76
Zones non végétalisées	0	40,37	0	0	0	0	0	0	1,11
Prairie	0	0	100	55,27	100	79,38	0	0	10,28
Prairie_arborée-arbres isolés	0	59,63	0	44,73	0	0	0	0	2,48
cultures	0	0	0	0	0	0	0	0	0
Pelouse	0	0	0	0	0	0	6,1	0	2,75
Maquis-garrigue-forêt_de_feuillus	0	0	0	0	0	20,62	89,7	4,81	42,81
Eau	0	0	0	0	0	0	0	95,19	29,82
unclassified	0	0	0	0	0	0	0	0	0
Total	100	100	100	100	100	100	100	100	100

Précision totale : 58,75 %

**Figure 10 : Matrice de confusion de la classification avec optimisation d'un espace d'attributs**



Quatre classes présentent d'importantes confusions. Les « zones non végétalisées » et les « prairies arborées » présentent respectivement des confusions avec les « prairies arborées » et les « prairies ». Les classes « cultures » et « pelouse » ont un taux nul de pixels bien classés. La classe « culture » ne compte que très peu de relevés (3 parcelles de vigne et une de maïs) car les oliveraies, qui représentent en terme de surface, la culture la plus importante dans ce secteur, a finalement été classée avec « Maquis, Garrigue, Forêt » à cause de confusions trop importantes. La précision n'atteignant pas les 60 %, cette classification est peu fiable.

Une classification basée uniquement sur les moyennes de valeurs des 4 canaux de l'image SPOT (PIR, rouge, vert et MIR) et sur le NDVI a ensuite été testée. Le résultat en est nettement plus satisfaisant comme le montre la matrice de confusion suivante.

Classif \ Vérité terrain	Pinède	Zones non végétalisées	Prairie	Prairie_arborée-arbres isolés	cultures	Pelouse	Maquis-garrigue-forêt_de_feuillus	Eau	Total
	Pinède	83,97	0	0,29	0,96	0	1,35	10	7,19
Zones non végétalisées	0	72,84	6,05	0	0	0	0	0,01	2,33
Prairie	0	0	65	39,46	0	0	0	0	4,13
Prairie_arborée-arbres	8	24,57	28,71	58,95	0	9,21	0,47	0,13	3,98
cultures	0	0	0	0	100	0	0	0	0,35
Pelouse	0,13	0,65	0	0	0	89,43	1,08	0	4,65
Maquis-garrigue-forêt_de_feuillus	15,63	1,95	0	0,64	0	0	88,45	1,28	41,64
Eau	0	0	0	0	0	0	0	91,4	28,63
unclassified	0	0	0	0	0	0	0	0	0
Total	100	100	100	100	100	100	100	100	100

**Précision totale : 81,3 %**

**Figure 11 : Matrice de confusion de la classification sur la radiométrie et NDVI – secteur Ajaccio**

Les confusions les plus importantes sont entre les classes « prairie » et « prairie arborée – arbres isolés » ce qui peut, à priori, paraître logique étant donné la nature proche de ces deux classes. Des confusions entre « zones non végétalisées » (ZNV) et « cultures » ont été remarquées visuellement (ces confusions ne sont pas représentées dans la matrice). Celles-ci s'expliquent par la structure des parcelles de vignes, qui alternent rang de vigne et sol nu, et qui ont donc une signature spectrale proche des sols nus (classés dans « ZNV »). La classe « prairie arborée – arbres isolés » se confond aussi avec les quartiers urbains, classés dans les « ZNV », dans lesquels on retrouve des jardins et des arbres épars expliquant la confusion. Le taux de précision totale est élevé, cette classification, présentée ci-après, est donc gardée pour la suite de l'étude.



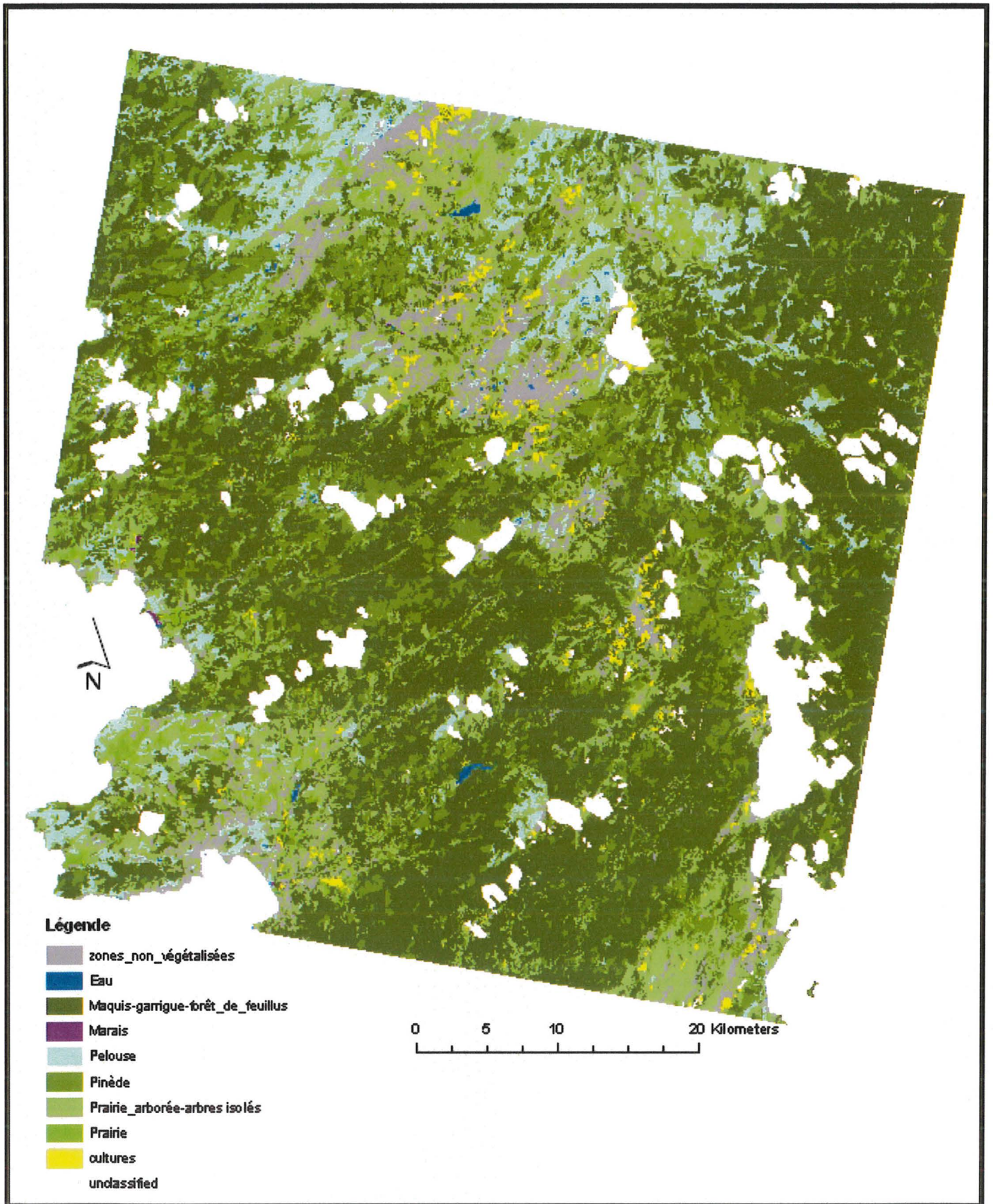


Figure 12 : Classification orientée objet de l'image SPOT – secteur Ajaccio



### 2.3.5.2. Classification de l'image SPOT – secteur Var

Selon le même protocole, une première classification a été testée avec définition automatisée d'un espace de 40 attributs via la fonction *feature space optimisation* (cf annexe 7). Là encore les résultats issus de la matrice de confusion sont peu satisfaisants.

V. terrain Classif.	Maquis	Forêt feuillus	Pinède	Vignoble	Prairie	Pelouse	Bâti sol_nu	routes principales	quartier urbain	Eau	zones brulees	Sum
Maquis	57,08	33,06	0	39,14	0	0	0	1,32	0	0	0	8,02
Forêt feuillus	0	0	0	0	0	0	0	0	0	0	0	0
Pinède	40,92	57	99,34	0	0	0	0	3,05	0	0	0	15,53
Vignoble	0	6,25	0	51,2	0	0	0	0,15	0	0	0	4,27
Prairie	9	0	0,66	0	32,68	0	0	20,79	0	0	0	3,34
Pelouse	0	3,7	0	0	67,32	0	0	0	0	0,1	0	2,16
Bâti sol_nu	0	0	0	0	0	0	100	0	25,58	0	0	1,6
Routes	0	0	0	0	0	0	0	57,26	0	0,51	0	6,96
quartier urbain	0	0	0	0	0	0	0	0	74,42	0	0	4,28
Eau	0	0	0	1,75	0	0	0	0	0	98,3	0	47,1
zones brulees	0	0	0	0	0	100	0	0	0	0	100	3,66
unclassified	0	0	0	7,84	0	0	0	17,44	0	1,05	0	3,08
Sum	100	100	100	100	100	100	100	100	100	100	100	100

**Précision totale : 60,94 %**

**Figure 13 : Matrice de confusion de la classification avec optimisation d'un espace d'attributs – secteur Var**

Aucun des objets des classes de validation « forêt feuillus » et « pelouse » n'est correctement classé. Les classes « maquis », « vignoble » et « prairie » présentent des confusions trop importantes pour être fiable.

Une classification basée sur la radiométrie et sur le NDVI a aussi été testée pour ce secteur, donnant une nouvelle fois de bons résultats, le taux de précision étant de 94,96 %.

	Foret feuillus	Routes principales	Eau	Zones_ brulees	Vignoble	Maquis	Prairie	Bâti sol_nu	Pinède	Urbain	Pelouse	Total
Foret_feuillus	95,62	0	0,02	0	0	1,99	0	0	1,13	0	0	11,58
Routes_principales	0	94,55	0,05	0	0	0	0	0	0	0	0	10,84
Eau	0	0	99,8	0	0	0	0	0	0	0	0	46,46
Zones_brulees	0	0	0,13	100	0	0	0	0	0	0	0	4,42
Vignoble	0	1,24	0	0	80,17	0	2,33	0	0	1,3	0	5,59
Maquis	4,38	0,71	0	0	0	98,01	0	0	13,07	0	0	3,45
Prairie	0	0	0	0	0	0	91,9	0	0	0	0	2,2
Bâti_sol_nu	0	0	0	0	0	0	0	100	0	0	0	0,12
Pinède	0	0	0	0	0	0	0	0	85,8	0	0	6,18
Urbain	0	2,26	0	0	18,6	0	4,67	0	0	98,7	0	7,13
Pelouse	0	1,24	0	0	1,23	0	1,08	0	0	0	100	2,02
unclassified	0	0	0	0	0	0	0	0	0	0	0	0
Total	100	100	100	100	100	100	100	100	100	100	100	100

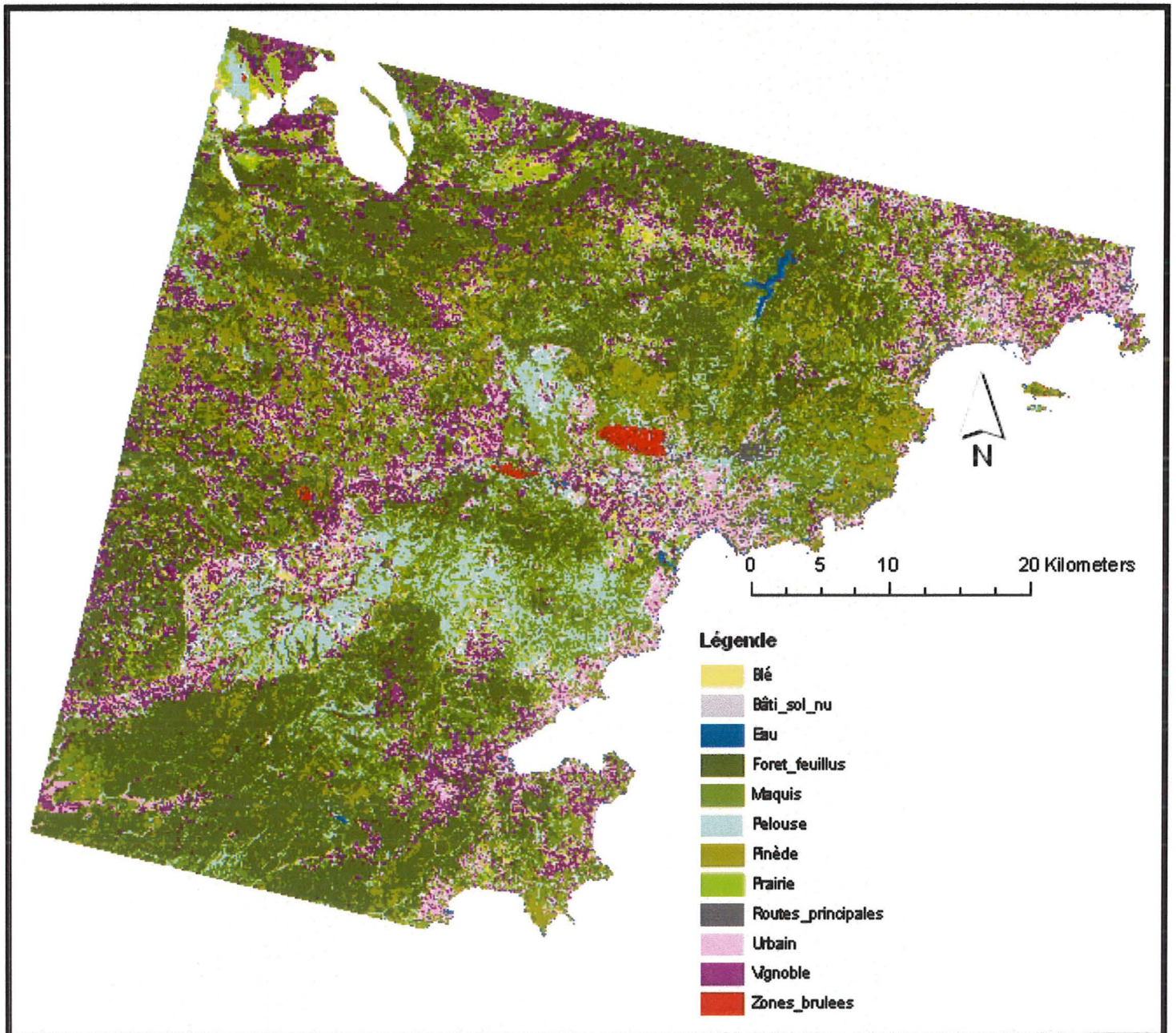
**Précision totale : 94,96 %**

**Figure 14 : Matrice de confusion de la classification sur la radiométrie et le NDVI**

Ce résultat ne signifie pas nécessairement que la classification est bien représentative de la réalité, mais simplement que les parcelles de validation ont été correctement classées. La



surface représentée par la totalité des relevés, pour la Corse et pour le Var, ne représente qu'une infime proportion de la surface totale correspondant aux images. De tels résultats auraient eus une valeur significative bien plus importante si une vérité terrain représentant une proportion non négligeable de la superficie de l'image avait été confrontée à ces classifications. L'interprétation visuelle confirme néanmoins ce bon résultat.



**Figure 15 : Classification orientée objet de l'image SPOT Var**

L'apparition d'une classe « zones brûlées » est à noter. Ces zones avaient été préalablement remarquées sur l'image en composition colorée ainsi que sur l'ACP et



identifiées lors de la mission. Les classes « maquis » et « forêt feuillus » sont dans des classes distinctes et différents types de « zones non végétalisées » ont été différenciées. Ces classes pourront par la suite être regroupées pour appliquer les mêmes variables qui ont servi à établir les modèles de risque. Malgré les bons résultats de la matrice, d'importantes confusions facilement identifiables visuellement apparaissent entre la classe « urbain » et la classe « vignoble ». Une grande partie des objets ainsi mal classés entre ces deux classes ont été reclassés manuellement.

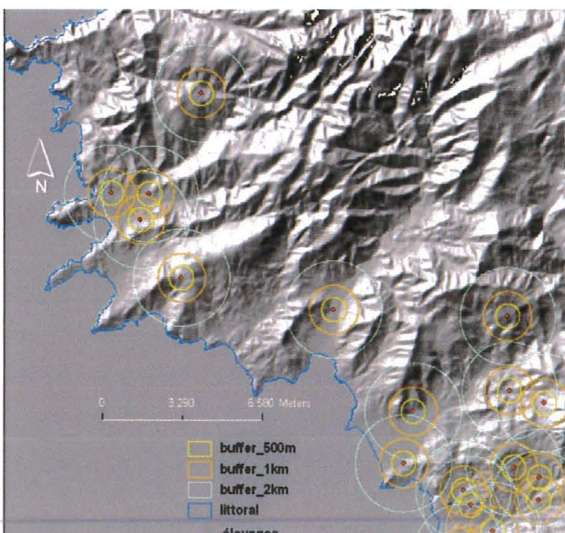
## 2.4. Extraction des variables environnementales

L'extraction des variables liées à la végétation et à la topographie a été réalisée, en intégrant les données dans un SIG (ArcGis® et Arcview®), à l'intérieur de trois zones tampons (500 mètres, 1 et 2 kilomètres), autour des élevages (Corse) et piégeages (Var).

L'extraction des variables paysagères issues des cartes d'occupation du sol s'est faite via le logiciel Fragstats.

La variable *Y* renseignant sur la latitude des élevages et des piégeages est issue des relevés GPS qui ont servi à leur localisation. Seul le modèle établi à 1 km n'inclut pas cette variable, qui est calculée en mètres pour le modèle établi à 500 m avec variable paysagère et à 2 km. L'autre modèle établi à 500 m inclut cette variable calculée en kilomètres. La latitude est une variable explicative importante étant donnée la provenance du moucheron et sa propagation (globalement du sud vers le nord : il est arrivé en Corse depuis la Sardaigne et ensuite dans Var).

### 2.4.1. Extraction issue du MNT : l'ensoleillement moyen



L'ensoleillement est le seul paramètre lié à la topographie retenu dans deux des quatre modèles statistiques (modèles établis à 1 et 2 km). Il est calculé à 16 heures le 30 juin.

C'est un indicateur de la température et de l'humidité du milieu.

**Figure 16 : Ensoleillement issu du MNT d'Ajaccio & zones tampons autour des élevages**

## 2.4.2. Extraction des pourcentages de classe d'occupation du sol.

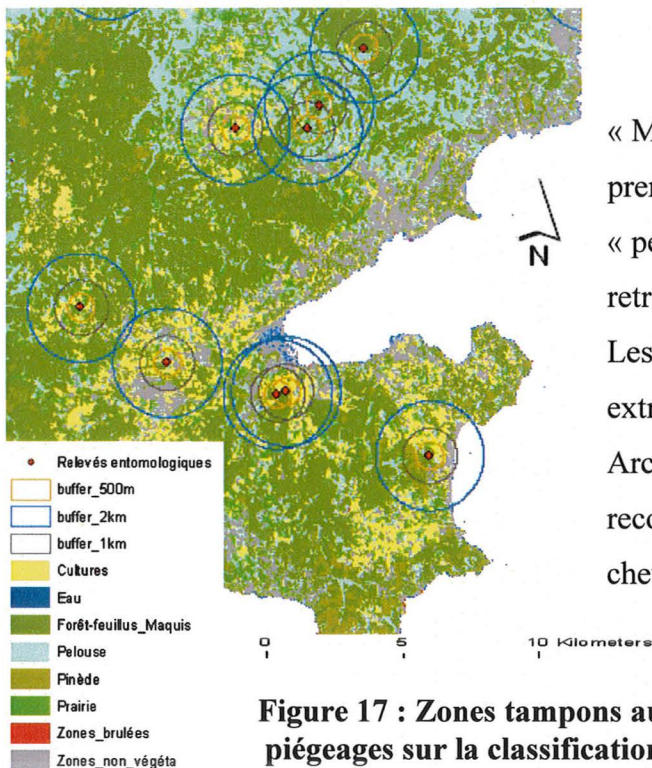


Figure 17 : Zones tampons autour des piégeages sur la classification du Var

Les pourcentages des classes « cultures » et « Maquis, Garrigue, Forêt » sont des variables du premier modèle établi à 500 mètres. La classe « pelouse », recodée en présence/absence se retrouve dans le modèle établi à 1km.

Les aires des classes d'occupation du sol ont été extraites via l'extension *Spatial analyst* dans Arcview®, qui permet de ne pas tenir compte du recouvrement des zones tampons qui se chevauchent.

## 2.4.3. Extraction des variables paysagères

### 2.4.3.1. Définition

Les variables paysagères traduisent la configuration du paysage au sein des zones tampons. Plus exactement, elles témoignent de la structure de répartition spatiale des différentes classes d'occupation du sol. Elles rendent compte des effets de lisière, de corridor, de continuité, de contiguïté. La recherche du rôle de ces variables part de l'hypothèse que l'organisation du paysage influe sur la répartition de *C. imicola*. Les différentes variables paysagères extraites pour ces modèles et pour la validation sont issues du logiciel Fragstats®. Celui-ci permet d'extraire de nombreuses informations sur trois niveaux différents : le paysage, les classes et les « patches ».

Le modèle établis à 500 m, 1 et 2 km ont respectivement comme variables paysagères :



- *LS PRD (Landscape Patch Richness Density<sup>25</sup>)* : il témoigne de la diversité des classes présentes à l'intérieur des zones tampons. Son unité de mesure est le nombre de classe par 100 hectares.
- *LSI (Landscape Shape Index<sup>26</sup>)* de « prairie » et « zones non végétalisées » : Le *LSI*, appliqué à une classe, est une mesure de l'agrégation et de la compacité de cette classe à l'intérieur des zones tampons. Il rend compte du nombre de bordures des objets d'une classe dans celles-ci.
- *NP (Number of Patches<sup>27</sup>)* « prairie » : C'est un indice de la fragmentation d'une classe dans les zones tampons. Seul le nombre de patch de la classe considérée est extrait.

2.4.3.2. Méthode d'extraction

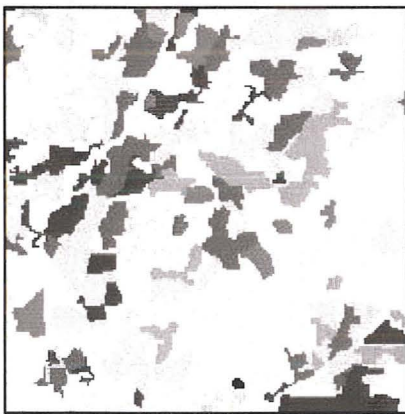


Figure 18 : exemple « d'imagette » extrait à l'échelle de 1 km

Ces variables ont pu être extraites grâce aux classifications réalisées sur des images à haute résolution spatiale (10 mètres) qui permettent cette analyse propre aux échelles fines.

Un programme<sup>28</sup> permettant d'extraire des « imagettes » autour des élevages et des piégeages sur les cartes d'occupation du sol (en format raster) a été créé (A. Tran) et appliqué sous Envi®.

Les images de sorties sont carrés car le programme extrait d'autant de pixel à l'horizontale, à la verticale qu'en diagonale. Cette forme a été gardée, une approche sur des images rondes ne changeant pas les valeurs des variables.

Ces images sont ensuite traitées dans Fragstats® qui permet d'extraire les types

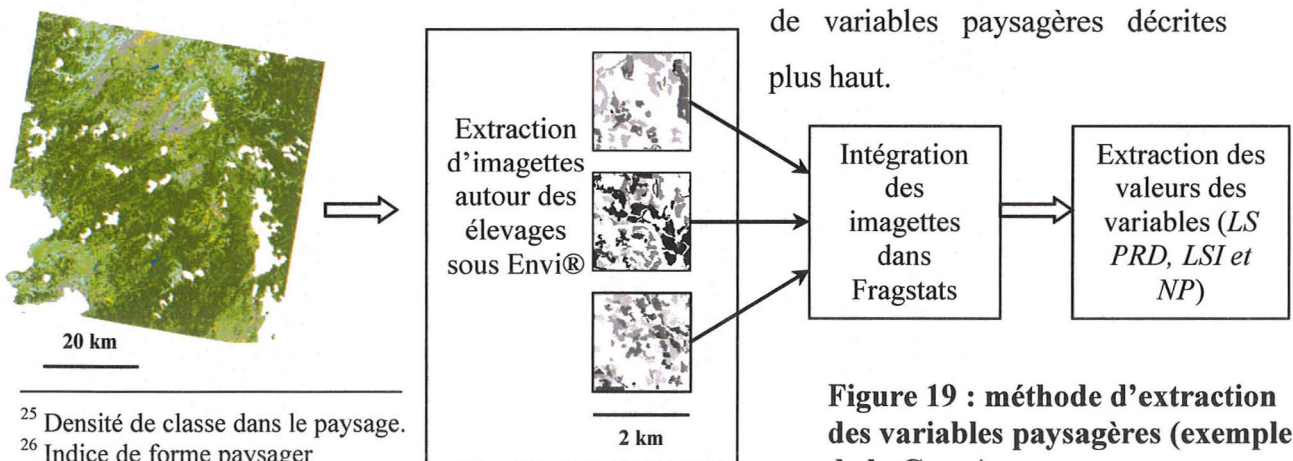


Figure 19 : méthode d'extraction des variables paysagères (exemple de la Corse)

<sup>25</sup> Densité de classe dans le paysage.  
<sup>26</sup> Indice de forme paysager  
<sup>27</sup> Nombre d'objet  
<sup>28</sup> Réalisé en IDL

### 3. Validation externe des modèles établis en Corse du Sud

#### 3.1. Calcul des probabilités de risque

L'ensemble des variables composant les 4 modèles ont donc été extraites et organisées dans un tableau Excel® afin de calculer respectivement pour chaque élevage et pour chaque piégeage les probabilités qu'il soit infecté ou positif en *C. imicola*. L'application de modèles établis à partir de données épidémiologiques au Var (donc sur des données entomologiques) nécessite l'explication de deux hypothèses :

- Les piégeages avérés positifs à la présence de *C. imicola* sont considérés comme des foyers infectés (et inversement) par la maladie pour la confrontation des prédictions à la réalité.
- L'ensemble des modèles incluant la variable *mixité*, deux scénarios par modèle ont été réalisés, considérant alternativement les piégeages comme *mixtes* et *non mixtes* pour le calcul des probabilités de risque.

Les variables sur les pourcentages d'occupation du sol et sur l'ensoleillement ont été recodées entre 0 et 1 et les variables paysagères, qui ne sont pas associées à un coefficient, recodées en tertiles<sup>29</sup>. Chaque élevage et piégeage est ainsi caractérisé par une probabilité de risque d'infection comprise entre 0 et 1, selon l'équation présentée en 1.2.2. La connaissance du statut sanitaire réel pour chaque élevage de Corse et de la présence ou non de *C. imicola* dans les piégeages du Var permet de confronter les prédictions des modèles à la réalité. L'application d'un seuil de 0,5 pour discriminer les prédits positifs et négatifs ne donnent pas de résultats, les valeurs des probabilités, pour les modèles établis à 500m et 2km étant largement et majoritairement inférieures à ce seuil. Cependant, un tri croissant des probabilités révèle pour certains modèles un bon classement, dans la mesure où l'on retrouve beaucoup d'élevages sains dans les valeurs de prédiction les plus basses et inversement. Ces modèles, appliqués tel quel en externe, ne peuvent être considérés en quantitatifs (seuil de 0,5 non représentatif de la discrimination entre les élevages sains et infectés) mais en qualitatifs dans la mesure où les probabilités constituent ici un indice de risque relatif. Par conséquent, la détermination d'un seuil adéquate est nécessaire dans l'optique d'une discrimination optimale des prédictions des statuts sanitaires des élevages.

---

<sup>29</sup> Recodage d'après le découpage en tertiles de la distribution des valeurs associées aux variables.



### 3.2. Détermination d'une valeur seuil et évaluation du pouvoir discriminant des modèles de risque : analyse de ROC

Evaluer la performance de l'application externe des modèles de risque passe par la détermination d'un seuil permettant la meilleure discrimination entre les élevages infectés et indemnes. L'analyse de ROC (*Receiver-operating characteristic*<sup>30</sup>) est un outil permettant à la fois d'évaluer la capacité d'un modèle à discriminer et de déterminer une valeur seuil de probabilité de risque permettant la meilleure discrimination. Cette analyse amène à considérer les notions de sensibilité (Se) et de spécificité (Sp). La première correspond à la probabilité que la prédiction soit positive chez les élevages infectés et se calcule par le ratio suivant :

$$Se = \text{Vrai Positifs} / (\text{Vrai Positifs} + \text{Faux Négatifs})$$

La spécificité correspond à la probabilité que la prédiction soit négative chez les élevages indemnes :

$$Sp = \text{Vrai Négatifs} / (\text{Vrai Négatifs} + \text{faux Positifs})$$

Les taux de sensibilité et de spécificité varient, par conséquent, en fonction du seuil arrêté. Il est donc possible, suivant les objectifs du modèle, de privilégier l'une ou l'autre notion : si l'objectif est le dépistage, la sensibilité est privilégiée. La spécificité est quand à elle-même privilégiée pour un diagnostic de certitude. La figure suivante illustre plus clairement ces propos.

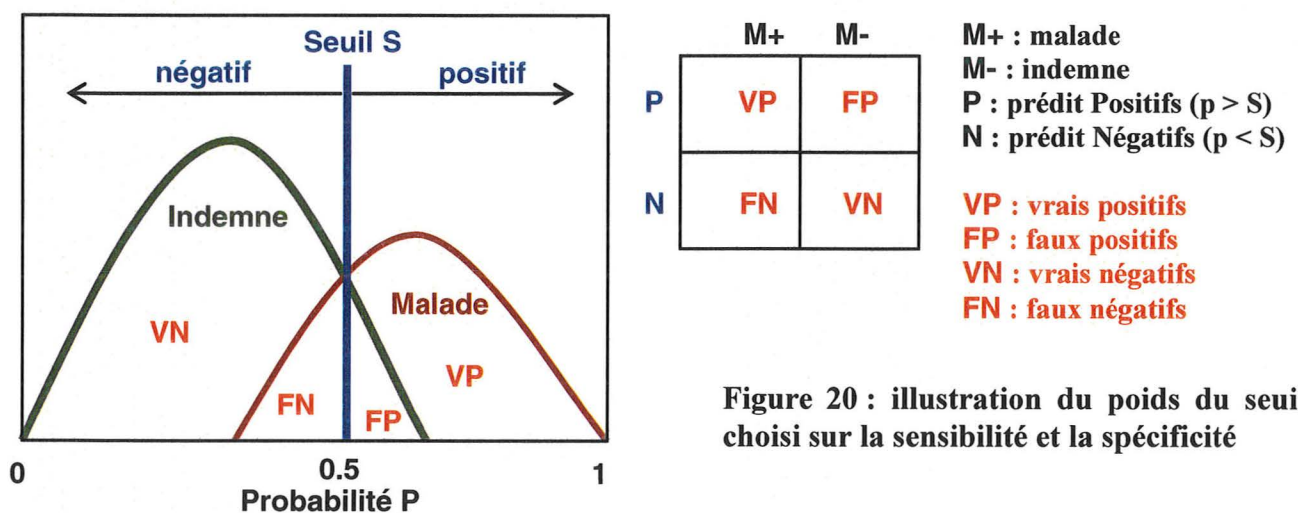
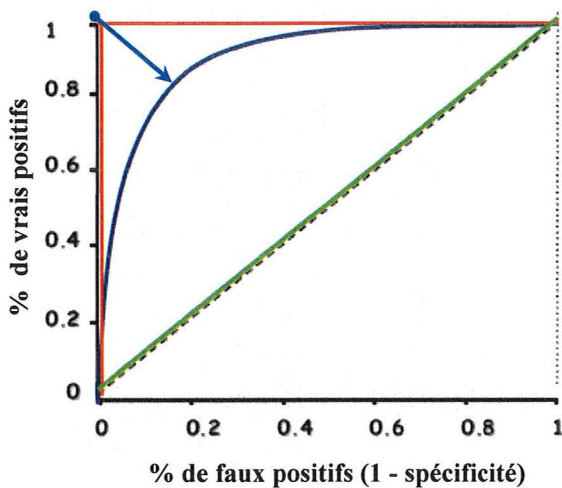


Figure 20 : illustration du poids du seuil choisi sur la sensibilité et la spécificité

L'objectif de ces modèles étant, à terme, de cibler la surveillance dans les zones où le vecteur est présent, la sensibilité est privilégiée. Ainsi, le seuil choisi sera plutôt abaissé de

<sup>30</sup> Littéralement, « Caractéristique de fonctionnement du récepteur »

manière à diminuer le nombre d'élevages infectés mais prédits indemnes (= FN). Ce choix se fait nécessairement au dépend de la spécificité.



**Figure 21 : Exemple de courbes de ROC**

L'aire sous la courbe est un indicateur du pouvoir discriminant du modèle. Elle varie entre 0 et 1 (0 étant la diagonale verte, 1 la courbe rouge passant par les coordonnées  $\{0;1\}$  et 0,8 la courbe bleue). Les modèles peuvent ainsi être classés en fonction de leur niveau de performance<sup>31</sup> :

- si aire ROC = 0,5 → il n'y a pas de discrimination.
- si aire  $0,7 \leq \text{ROC} < 0,8$  → la discrimination est acceptable.
- si aire  $0,8 \leq \text{ROC} < 0,9$  → la discrimination est excellente.
- si aire ROC  $\Rightarrow 0,9$  → la discrimination est exceptionnelle.
- si aire ROC = 1 → la discrimination est parfaite

Le seuil optimum pour obtenir un taux maximum d'élevages correctement classés correspond à la probabilité qui permet, en prenant l'exemple de la courbe bleue, de minimiser la distance (flèche bleue) au point idéal  $\{0;1\}$ , où sensibilité = spécificité = 1.

Le logiciel STATA® a été utilisé pour la réalisation des courbes de ROC. Les probabilités pour chaque élevage ainsi que leurs statuts réels sont intégrés en entrées. L'aire sous la courbe est automatiquement donnée ainsi qu'un tableau avec le détail sur la sensibilité et la spécificité pour chaque probabilité de risque prédite. Le taux de foyer correctement classé y est également indiqué pour chaque valeur seuil. Il est ainsi possible de choisir le seuil en fonction des objectifs, soit en privilégiant la sensibilité (avec un taux supérieur ou égal à 70 %) tout en prenant soin de conserver un taux d'élevage bien classé le plus important possible. Les seuils choisis correspondent au meilleur compromis entre un taux d'élevage bien classé et un bonne sensibilité.

<sup>31</sup> Le classement présenté ici est issu du *Cours de régression logistique appliquée*. Taffé P. Institut Universitaire de Médecine Sociale et Préventive (IUMSP) et Centre d'épidémiologie Clinique (CepiC). Lausanne, août 2004.

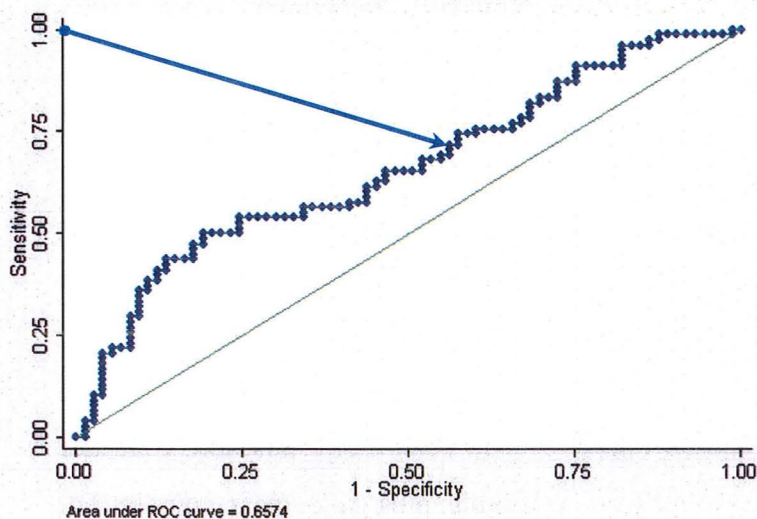


### 3.3. Résultats de l'application externe des modèles

#### 3.3.1. Validation dans le secteur d'Ajaccio

##### 3.3.1.1. Modèles établis à 500 mètres :

- Modèle sans variable paysagère :

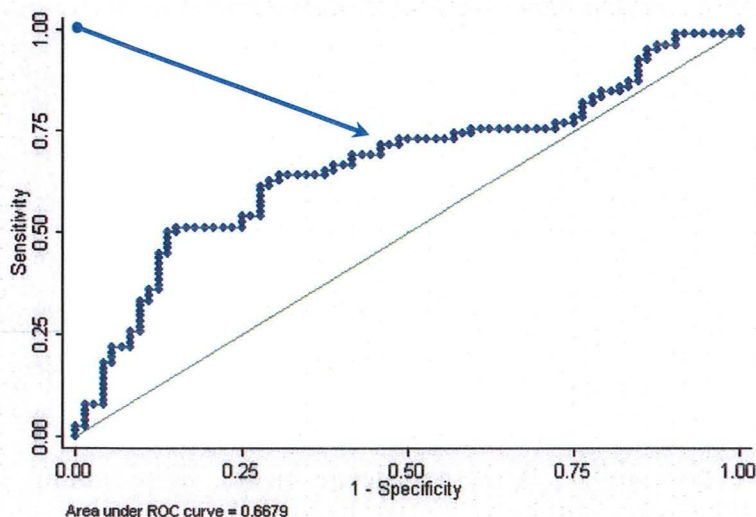


L'aire sous la courbe (AUC) est de 0,6574, ce qui révèle une discrimination relativement faible du modèle. Le taux maximum d'élevage correctement classé est de 64,90 %, avec un seuil de 0,089775 où la sensibilité est de 50 % et la spécificité à 80,82 %. Un seuil défini subjectivement en fonction de l'objectif de privilégier la sensibilité, à 0,012916 (flèche bleue) donne une sensibilité de 74,36 % mais avec une spécificité

**Figure 22 : Courbe de ROC du modèle établis à 500 m sans variable paysagère**

égale à 42,47 %. Le taux maximum d'élevage bien classé est alors de 58,94 %. L'application externe de ce modèle n'offre donc pas un résultat satisfaisant.

- Modèle avec variable paysagère :

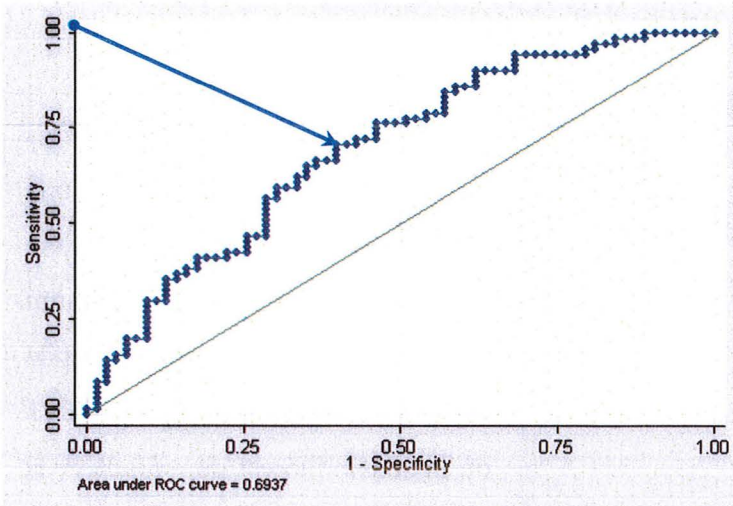


Le résultat est ici proche que pour le modèle précédent. L'AUC est de 0,6679 avec un taux maximum d'élevage correctement classé à 67,33 % (seuil de 0,003628 ; sensibilité = 51,28 % et spécificité = 84,72 %). La sensibilité étant privilégiée, un seuil fixé à 0,000749 (flèche

**Figure 23 : Courbe de ROC du modèle établis à 500 m avec variable paysagère**

bleue) offre une sensibilité de 71,79 % et une spécificité de 54,17 % pour un taux d'élevage bien classé proche du taux maximum (63,33 %). Ce modèle présente ainsi une amélioration notable par rapport au précédent mais reste encore insuffisamment discriminant.

**3.3.1.2. Modèle établi à 1km :**

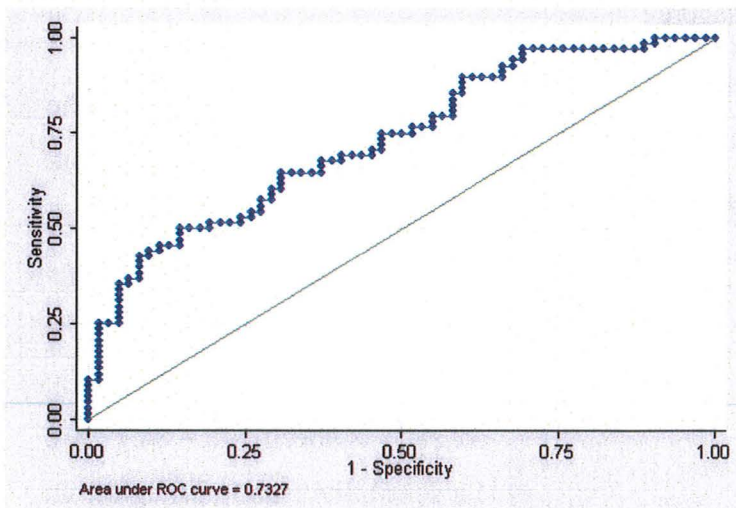


Avec une AUC proche de 0,7, ce modèle peut être considéré comme acceptable. Le taux d'élevage correctement classé atteint son maximum avec un seuil fixé à 0,039332 (flèche bleue). La sensibilité est alors de 70,42 % et la spécificité de 60,32 %, ce qui dénote une discrimination satisfaisante. Le taux de sensibilité peut encore être amélioré avec des seuils plus bas, mais cela se fait

**Figure 24 : Courbe de ROC du modèle établi à 1 km**

alors au dépend de la spécificité qui ne doit pas être totalement négligée même si l'objectif n'est pas de privilégier. Le seuil énoncé répond aux exigences fixées, le taux de bien classer étant optimal et la sensibilité supérieure à 70 %.

**3.3.1.3. Modèle établi à 2 km :**



Pour le secteur d'Ajaccio, ce modèle a le plus fort pouvoir discriminant. L'AUC est de 0,7327, ce qui rend ce modèle acceptable. Le taux maximal de bien classé atteint 66,92 % avec un seuil de 0,000515. La sensibilité et la spécificité sont respectivement de 64,71 % et de 69,35 %. Si le seuil est abaissé à 0,000177 (flèche bleue), la sensibilité

**Figure 25 : Courbe de ROC du modèle établi à 2 km**

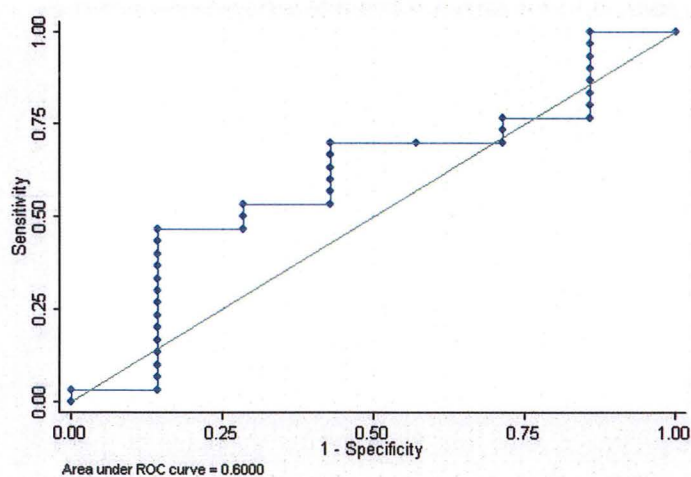


s'élève à 75 % et la spécificité tombe à 53,23 % pour un taux de bien classé de 64,62 %. Ce modèle est donc celui qui s'exporte le mieux sur ce secteur de la Corse

### 3.3.2. Validation dans le secteur du Var

L'application des modèles de risque dans ce secteur n'a pas été concluante. Tous les modèles incluant la variable  $Y$  ont donné des probabilités trop faibles (de l'ordre de  $10^{-28}$  pour les valeurs les plus élevées) pour permettre de réaliser des courbes de ROC. Seul le modèle établi à 1 km, qui n'inclus pas  $Y$  donne des probabilités dans une échelle de grandeur exploitable (entre 0,0018 et 0,99).

#### 3.3.2.1. Modèle établi à 1km :



**Figure 26 : Courbe de ROC du modèle établi à 1 km**

(notamment pour la variable  $Y$ ).

Les deux courbes, pour les scénarios où les piégeages sont considérés respectivement comme mixtes, puis non mixtes ont donné les mêmes résultats. L'AUC est de 0,6, le modèle est donc ici très peu discriminant. Il semble donc que l'essai de validation externe sur le continent, et qui plus est sur des données entomologiques ne puisse se faire sans adapter les modèles de prédiction

## 4. Discussion et perspectives

### 4.1. Discussion sur les données relatives à la bluetongue

#### □ Les données épidémiologiques

Utilisées pour la validation dans le secteur d'Ajaccio, les données fournies pour ce travail semblaient cependant ne pas être exhaustives. Le recueillement des caractéristiques propres à chaque élevage sur une si grande zone d'étude depuis 2001 est extrêmement complexe. Il en

ressort des biais éventuels sur la validation des modèles qui intègrent deux variables directement liées à ces données (sauf le modèle établi à 1km qui n'en intègre qu'une) :

- *Y* : les élevages sont géoréférencés au niveau de la bergerie. Un travail sur la localisation des pâtures aurait été extrêmement complexe et lourd à mettre en place. Cependant la transmission du virus par *C. imicola* est nocturne et les élevages sont généralement rentrés à la bergerie pour la nuit ce qui conforte l'étude à partir de données localisées à cet endroit. Par contre, le paramètre de transhumance propre aux élevages n'est pas inclus dans cette analyse alors qu'il pourrait être un facteur explicatif de la transmission. Là encore, ces données sont extrêmement complexes à recueillir et seraient d'autant plus difficiles à exploiter.
- *La mixité* : cette information fait partie des données parfois manquantes dans le recueil des caractéristiques des élevages. Cela fausse directement le calcul de probabilité de risque de certains élevages et l'évaluation finale des modèles en externe.

Les données sur l'état sanitaire des ovins semblent fiables et relativement exhaustives, la maladie étant à déclaration obligatoire, accompagnée d'une indemnisation.

#### □ **Les données entomologiques :**

Relatives au secteur du Var, ces données ont servi pour un essai de validation des modèles « maladies » sur des données vectorielles. Peu nombreuses pour cette étude (38 pièges ont pu servir à la validation) car les campagnes de piégeage sont lourdes à mettre en place et que le vecteur vient de s'y installer, deux nouvelles campagnes en juin et en septembre 2006 vont permettre d'étoffer la base de données de répartition du moucheron.

Il faut aussi souligner que pour la validation des modèles, seules les données épidémiologiques et entomologiques dont les zones tampons recouvrent au minimum 75 % des images ont été gardées de façon à ne pas biaiser le calcul des probabilités de risque.

#### □ **Perspectives :**

Les pièges sont quasiment tous positifs dans le Var mais avec des nombres d'individus capturés très variables. La validation réalisée dans ce rapport était simplement qualitative, puisque basée sur la présence ou non de *C. imicola* dans les pièges. Une étude quantitative semblerait donc plus pertinente au regard de la disproportion dans la répartition du vecteur.



La variable  $Y$  pourrait aussi être changée, pour ce secteur, en distance par rapport à la Corse. En effet, cette variable a été incluse dans les modèles à partir d'une étude en Corse du sud, le vecteur arrivant directement de Sardaigne dans un axe nord/sud. Cependant, d'après les modèles, plus les élevages sont au nord moins le risque d'infection est important. Cela explique que l'application dans le Var (modèles à 500m et 2km) donne des probabilités extrêmement faibles. Le modèle à 1 km n'incluant pas le  $Y$  donne des probabilités dont la distribution paraît plus « normale » de 0,2 à 0,99 (sur une échelle de 1). La distance par rapport à la Corse pourrait éventuellement être plus représentative de l'explication de la présence du moucheron sur le continent.

## **4.2. Discussion sur les paramètres environnementaux**

### **□ Méthode de traitement des images :**

Les bons résultats des matrices de confusions des classifications réalisées en Corse et dans le Var ont été discutés à la suite de leurs présentations dans le § 2.3.5.

Le principe de la classification par une approche orientée objet prend aussi son importance dans l'extraction des variables paysagères. En effet, la phase de segmentation de l'image en objets homogènes conditionne à priori les valeurs des variables paysagères qui sont calculées en fonction du nombre d'objet, de leur densité et de leur bordure dans les zones tampons. Le paramètre d'échelle prend ici toute son importance puisqu'il influe sur la structure représentée du paysage.

## **4.3. Discussion sur les modèles statistiques de risque**

Les validations des modèles établis à 1 et 2 kilomètres dans le secteur d'Ajaccio ont été relativement significatives, offrant un pouvoir discriminant assez fort. L'application de ces deux modèles dans ce secteur est par conséquent acceptable.

Les validations dans le Var ont fait ressortir d'importants problèmes d'adaptabilité des modèles. L'application sur des données entomologiques de modèles établis sur des données épidémiologiques ne semble pas pouvoir se faire sans certaines adaptations. Cet essai ne constituait qu'une première approche dans un secteur nouvellement touché par la maladie où les perspectives de travail sont importantes puisque la plupart des études ne sont actuellement qu'en cours.

□ **Perspectives :**

Ces résultats permettent d'envisager la réalisation d'une cartographie du risque de survenue de la maladie sur l'ensemble de la Corse. Cette dernière permettrait de cibler les zones à surveiller en priorité dans cette région particulièrement touchée par la fièvre catarrhale ovine.

Pour le Var, il serait intéressant de changer la variable  $Y$  en variable *distance Var / Corse* qui permettrait certainement d'obtenir des probabilités de risque plus forte. Les modèles appliqués prédisent en effet que plus on monte en latitude, moins le risque d'infection est élevé. De plus, l'application dans ce secteur nécessiterait certainement de recalculer les coefficients des modèles.



## Conclusion

Ce travail de validation a permis de mettre en évidence à la fois la robustesse des modèles de risque (établis à 1 et 2 km) dans un secteur adjacent à celui où ils ont été élaborés, ainsi que les limites de leur application dans un secteur plus éloigné et qui plus est sur des données de natures différentes. Les données de télédétection et les outils SIG, associés à un travail interdisciplinaire, sont des moyens de surveillance dans les thématiques épidémiologiques extrêmement puissants et difficilement incontournables depuis plus d'une décennie. Les perspectives issues des ces résultats sont importantes puisqu'elles permettent d'envisager l'application de ces modèles pour cibler efficacement les zones à surveiller en Corse et de cerner certains paramètres à modifier ( $Y$ ) et à réadapter (coefficients) pour l'exportation des ces modèles en métropole. De plus, des modèles de risques élaborés à partir de données entomologiques sont actuellement en cours [H. Guis] dans le même projet de thèse que celui dont on a validé les modèles. Il sera particulièrement intéressant de comparer ces nouveaux modèles à ceux déjà réalisés avec des données épidémiologiques.

Au delà des résultats obtenus et des perspectives envisagées, ce travail de stage, nécessitant des approches interdisciplinaires, nous a permis de développer une certaine rigueur scientifique et de nous enrichir dans de nombreux domaines de compétences : la télédétection par une approche orientée objet mais aussi de façon générale, les SIG, l'épidémiologie (puisque totalement néophyte en la matière), la gestion du travail sur le terrain, les statistiques, la botanique, le travail concerté, la gestion d'une importante base de données et des aléas informatiques.

## Bibliographie

Ancelle T. *Statistique épidémiologie*. Paris : Maloine, 2002. 300p

Baldet T., Mathieu B., Delécolle J.C, Gerbier G., Roger F. – Emergence de la fièvre catarrhale ovine dans le bassin méditerranéen et surveillance entomologique en France. *Revue Elev. Méd. Vét. Pays trop.*, 2005, **58 (3)**, 125-132.

Baldet T., Mathieu B., Delécolle J.C. – Emergence de la fièvre catarrhale ovine et surveillance entomologique en France. *Insecte*, 2005, n° 131, **4**, 28-30.

Benz U.C, Hofmann P., Willhauck G., Lingenfelder I., Heynen M. – Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISRPS Journal of Photogrammetry & Remote Sensing*, 2004, **58**, 239-258.

Burel F., Baudry J. *Ecologie du paysage : concepts, methods et applications*. Paris : TEC&DOC, 2003. 359p

De La Rocque S., Michel V., Plazanet D., Pin R. – Remote sensing and epidemiology : examples of applications for two vector-borne diseases. *Comparative Immunology, Microbiology & Infectious Diseases*, **27**, 2004, 331-341.

Galangau J., Loisel R. et participation de Barry JP. *Végétation du Var : les principales series de la végétation : carte schématique*. IMG, CRDPN : 1989. 64p

Gamisans J. *La végétation de la Corse*. Edisud : Aix en Provence, 1999. 391p

Garnier M., Delamare V., Delamare J., Delamare T. *Dictionnaire des termes de médecine*. Paris : Maloine, 1999. 973p

Girard M.C et C.M. *Traitement des données de télédétection*. Paris : Dunod, 1999. 529p

Guis H. – *Caractérisation par télédétection de l'environnement favorable aux culicoides vecteurs de la fièvre catarrhale ovine dans le bassin méditerranéen : cas de la Corse* ; sous la dir. de De Roger F. et De La Rocque S. Montpellier, Mastère SILAT : CIRAD, 2004. 54p.

Gardner I., Greiner M., Georgiadis M. – *Advanced Methods in Diagnostic Test : Evaluation and interpretation*. ISVEE 9, Breckenridge, Colorado, 2000. 107p

Hay S.I. – Remote sensing and disease control : past, present and future. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1997 , **91**, 105-106.

Perez Correa M. – *Développement d'une méthode de cartographie de l'occupation du sol le long des cours d'eau à partir de données de télédétection* ; sous la dir. de Wasson J.G. Mastère SILAT : CEMAGREF, 2004. 39p



Purse V.B, Mellor P.S, Rogers D.J, Samuel A.R, Mertens P.C, Baylis M. – Climate change and the recent emergence of bluetongue in Europe. *Nature Reviews*, 2005, **3**, 171-181.

Quertier P., Aboucaya A., Beltra S., Childeric M. *Guide du naturaliste dans le Var*. Seyssinet-Pariset : Libris, 2002. 382p

Rodhain F., Perez C. *Précis d'entomologie médicale et vétérinaire*. Paris : Maloine, 1985. 458p

Taffé P. *Cours de régression logistique Appliqué*. Institut universitaire de Médecine Sociale et Préventive (IUMSP) et Centre d'épidémiologie Clinique (CepiC), Lausanne, 2004. 59p

Tatem A.J, Baylis M., Mellor P.S, Purse B.V, Capela R., Pena I., Rogers D.J – Prediction of bluetongue vector distribution in Europe and north Africa using satellite imagery. *Veterinary Microbiology*, **97**, (2003), 13-29.

Tran A., Biteau-Coroller F., Guis H., Roger F. – Modélisation des maladies vectorielles. *Epidémiol. Et santé anim.*, 2005, **47**, 35-51.

Ward M.P. – Climatic factors associated with the infection of herds of cattle with bluetongue viruses. *Vet Res Commun*, 1996, **20**, 273-283.

Wittmann E.J., Mellor P.S, Baylis M. – Using climate data to map the potential distribution of *Culicoides imicola* (Diptera : Ceratopogonidae) in Europe. *Rev Sci Tech*, 2001, **20**, 731-740.

Zientara S., De La Rocque S., Gourreau J.M., Grégory M., Diallo A., Hendrikx P., Libeau Geneviève, Sailleau C., Delécolle J.C - La fièvre catarrhale ovine en Corse en 2000. *Epidémiologie et santé anim.*, 2000, **38**, 133-144.

## Tables des figures et tableaux

Figure 1 : progression de la FCO dans le bassin méditerranéen [Purse <i>et al.</i> , 2005].....	4
Figure 2 : Localisation des élevages de Corse du sud sur les images satellitales SPOT.....	11
Figure 3 : Localisation des élevages dans le Var.....	12
Figure 4 : Schématisation de la démarche de l'étude.....	12
Figure 5 : Schéma de la méthode de cartographie de l'occupation du sol.....	14
Tableau 1 : Caractéristiques des images SPOT 5.....	15
Figure 6 : Exemple de schéma permettant de positionner les relevés de végétation.....	15
Figure 7 : Interface de paramétrage de la segmentation de l'image dans eCognition®.....	17
Figure 8 : Illustration des contours des objets après segmentation (zoom de l'image SPOT 5 secteur Ajaccio).....	18
Figure 9 : Illustration des contours des objets après segmentation (zoom de l'image SPOT 5 secteur Var).....	19
Figure 10 : Matrice de confusion de la classification avec optimisation d'un espace d'attributs.....	21
Figure 11 : Matrice de confusion de la classification sur la radiométrie et NDVI – secteur Ajaccio.....	22
Figure 12 : Classification orientée objet de l'image SPOT – secteur Ajaccio.....	23
Figure 13 : Matrice de confusion de la classification avec optimisation d'un espace d'attributs – secteur Var.....	24
Figure 14 : Matrice de confusion de la classification sur la radiométrie et le NDVI.....	24
Figure 15 : Classification orientée objet de l'image SPOT Var.....	25
Figure 16 : Ensoleillement issu du MNT d'Ajaccio & zones tampons autour des élevages...26	
Figure 17 : Zones tampons autour des piègeages sur la classification du Var.....	27
Figure 18 : exemple « d'imagette » extrait à l'échelle de 1 km.....	28
Figure 19 : méthode d'extraction des variables paysagères (exemple de la Corse).....	28
Figure 20 : illustration du poids du seuil choisi sur la sensibilité et la spécificité.....	30
Figure 21 : Exemple de courbes de ROC.....	31
Figure 22 : Courbe de ROC du modèle établis à 500 m sans variable paysagère.....	32
Figure 23 : Courbe de ROC du modèle établis à 500 m avec variable paysagère.....	32
Figure 24 : Courbe de ROC du modèle établi à 1 km.....	33
Figure 25 : Courbe de ROC du modèle établi à 2 km.....	33
Figure 26 : Courbe de ROC du modèle établi à 1 km.....	34



# *Annexes*

**Annexe 1 : La maladie de la langue bleue doit son nom à une cyanose de la langue**



*Cliché P. Hendrikx – DSV 30*

**Annexe 2 : *Culicoides imicola***



**2 mm**

---

**Annexe 3 : Piège lumineux UV standard utilisé pour la capture des *Culicoides***



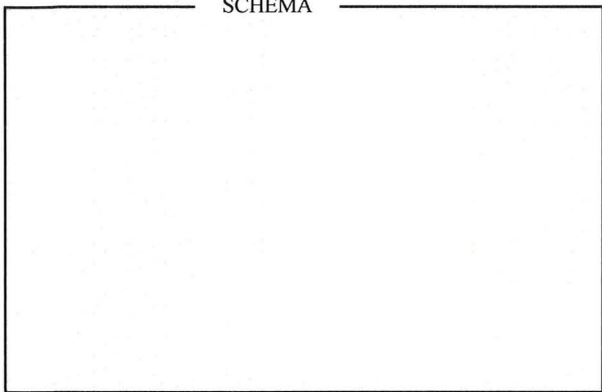


### Annexe 4 : Fiche de relevés de végétation

Site n° : .....  
 Secteur : .....  
 Date : ...../...../2006  
 Heure : .....  
 N° photo(s) terrain : .....

**Points GPS**

N°	X (m) (long)	Y (m) (lat)	Altitude



Remarques :

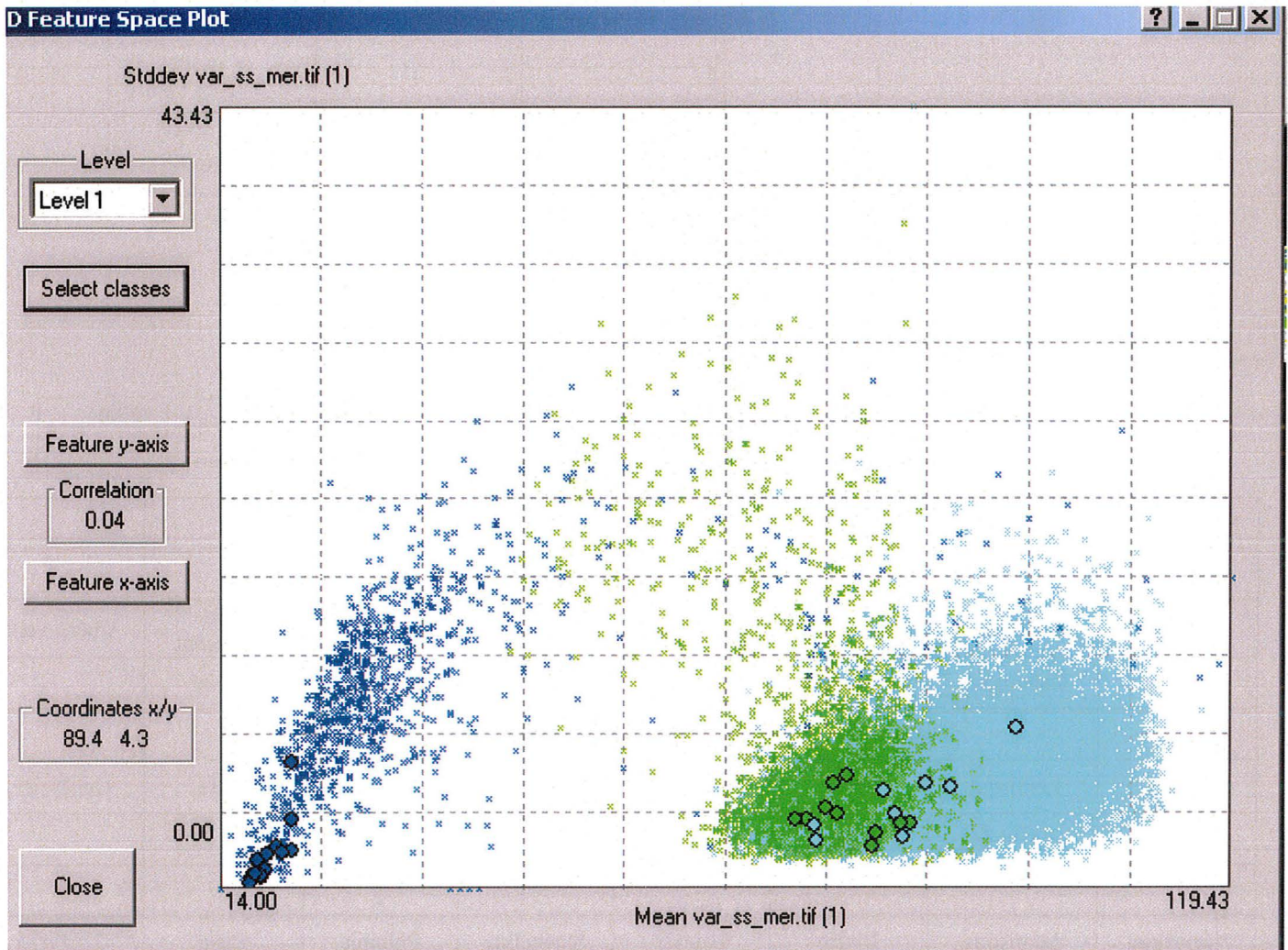
**Géomorphologie :**

Thalweg	Plaine	Plateau	Versant	Falaise	Rive	Vallée	Crête	colline

**Occupation du sol :**


Non végétal	Bâti - urbain	Sol nu	Sable Route	Roche			
<b>Forêt</b>	Pin d'alep	Pin maritime	Pin parasol	Pin sylvestre	Chêne liège	Chêne pubescent	Chêne vert
	Châtaigner	Charme	Hêtre	Aulne	Peuplier	Olivier	
<b>Maquis</b>	Lentisque	Myrte	Laurier	Stacheline	Phlomis	Ciste	Genévrier
	Lavande	Buis					
<b>Garrigue</b>	Lentisque	Myrte	Laurier	Stacheline	Phlomis	Ciste	Genévrier
	Lavande	Buis					
<b>Pelouse</b>	Lentisque	Myrte	Laurier	Stacheline	Phlomis	Ciste	Genévrier
	Lavande	Buis					
<b>Culture</b>	Vignoble	Oliveraie	Céréale	Terrasse à fleurs			
<b>Prairie</b>	Prairie rase	Prairie améliorée					
<b>Aquatique</b>	Marais	Ripisylve Aulne-	Rivière				

**Annexe 5 : Outil 2D Feature Space Plot : exemple de la distribution spatiale des classes d'occupation du sol (eau, pinède et pelouse)**





## Annexe 6 : attributs retenus pour la classification de l'image d'Ajaccio par l'outil *Feature Space Optimisation*

-  **Standard nearest neighbor (generated)**
  - GLCM Entropy (all dir.)
  - Mean diff. to scene Layer 4
  - Max. pixel value Layer 5
  - Mean of outer border Layer 2
  - GLCM Correlation Layer 1 (all dir.)
  - GLCM Correlation Layer 3 (all dir.)
  - GLCM Correlation Layer 5 (all dir.)
  - GLDV Entropy Layer 4 (all dir.)
  - GLCM Entropy Layer 1 (all dir.)
  - GLCM Entropy Layer 5 (all dir.)
  - GLCM Homogeneity (all dir.)
  - GLCM Homogeneity Layer 1 (all dir.)
  - GLCM Homogeneity Layer 3 (all dir.)
  - GLCM Correlation (all dir.)

## Annexe 7 : 40 attributs retenus pour la classification de l'image du Var par l'outil *Feature Space Optimisation*

Standard nearest neighbor (generated)	
Mean var_ss_nuage.tif (1)	
Ratio var_ss_nuage.tif (1)	
Mean diff. to scene var_ss_nuage.tif (1)	
Ratio to scene var_ss_nuage.tif (1)	
GLCM Correlation (all dir.), var_ss_nuage.tif (1)	
Min. pixel value var_ss_nuage.tif (2)	
GLCM Correlation (all dir.), var_ss_nuage.tif (2)	
StdDev. to neighbour pixels var_ss_nuage.tif (3) (0) (PxI)	
Min. pixel value var_ss_nuage.tif (3)	
GLCM Entropy (all dir.), var_ss_nuage.tif (3)	
GLCM Correlation (all dir.), var_ss_nuage.tif (3)	
Stddev var_ss_nuage.tif (4)	
StdDev. to neighbour pixels var_ss_nuage.tif (4) (0) (PxI)	
GLDV Entropy (all dir.), var_ss_nuage.tif (4)	
StdDev. to neighbour pixels bruno_ndvi.tif (0) (PxI)	
Max. pixel value bruno_ndvi.tif	
Length	
Width	
Compactness	
Elliptic. Fit.	
Rectangular Fit	
Border length	
Shape index	
Density	
Main direction	
Asymmetry	
Y center	
X min	
X max	
Distance to image border	
Perimeter (polygon)	
Compactness (polygon)	
Number of edges (polygon)	
Stddev of length of edges (polygon)	
Average length of edges (polygon)	
Length of longest edge (polygon)	
GLCM StdDev (all dir.)	
GLDV Entropy (all dir.)	
Number of right angles with edges longer than [1]	
Distance to line [1]	



