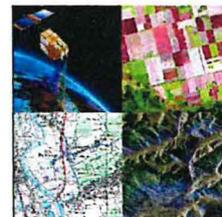
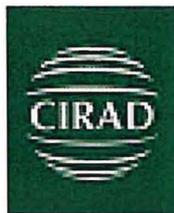


DK 54-1596

BA-TH 1476

G
2007



Utilisation des outils géomatiques pour la description d'un système infectieux multi hôtes

Cas du virus de West Nile dans le delta du Danube
(Roumanie)

Par
Arnaud Roux

Encadrant : Annelise TRAN
Tuteur : Sylvain THERY

CIRAD-Dist
UNITÉ BIBLIOTHÈQUE
Baillarguet

Stage de fin d'étude du 01/04/07 au 30/08/07
Master 2 Télédétection et Géomatique Appliquées à l'Environnement (Paris 7)



Remerciement :

Je tiens tout particulièrement à remercier :

- Annelise Tran, ma responsable, pour la confiance qu'elle m'a accordée et sa volonté de m'impliquer dans le projet global (réunion du projet EDEN à Paris, mission terrain en Camargue avec les entomologistes du projet, mission terrain en Roumanie et participation aux réunions West Nile et de l'UPR)
- Hélène Guis et Valéria Soti, membres de mon unité de recherche, pour leur soutien technique tout au long du stage
- Thierry Tomos et Stéphane Dupuy pour leur aide technique sur le logiciel eCognition
- Mihai Marinov, ornithologue du DDNI*, et principale personne impliquée dans mon travail, avec qui j'ai correspondu pendant toute la durée du stage. Je tiens à le remercier pour sa disponibilité pendant la période de mission et pour m'avoir servi de guide lors de la mission dans le delta
- Botond Kiss, responsable de l'équipe ornithologique du DDNI*, pour son accueil, sa disponibilité, sa volonté de gérer le séjour pour que je puisse réunir un maximum d'information
- Mihai Doroftei, botaniste du DDNI*, qui m'a accompagné pendant toute la partie terrain sur la zone sèche. Je tiens à le remercier pour m'avoir servi de guide
- Vasile, ornithologue en formation du DDNI*, pour nous avoir servi de chauffeur au cours des différentes missions
- Florian-Liviu Prioteasa, entomologiste au National Institute of Reserach and Developpement for Microbiology and Immunology « Cantacuzino », pour m'avoir réceptionné à Bucarest et m'avoir présenté ses zones de captures de moustiques
- Stefan Răileanu, ornithologue au DDNI*, pour son accueil et sa joie de vivre
- Toutes les personnes de mon unité de recherche qui m'ont accueilli chaleureusement au sein de l'équipe
- Toutes les personnes de la maison de la télédétection avec qui j'ai eu l'occasion de discuter

* DDNI : Danube Delta National Institute, laboratoire basé à Tulcea auquel j'ai été rattaché durant ma mission terrain en Roumanie

Table des matières :

Remerciement :	2
Introduction :	4
I) Présentation.....	5
1) Le cadre d'accueil	5
a) Le CIRAD	5
b) L'unité Epidémiologie et écologie des maladies animales	5
c) La Maison de la Télédétection.....	6
d) Le projet EDEN	6
2) Le virus de West Nile	7
a) Historique.....	7
b) Système épidémiologique	8
3) Contexte du stage	9
4) La zone d'étude.....	10
II) Réalisation d'une carte des habitats oiseaux.....	11
1) Les données utilisées.....	11
a) Carte de Végétation.....	11
b) Image Landsat	12
c) Modèle Numérique de Terrain.....	12
d) Données collectées sur le terrain.....	13
e) Carte des rivières et canaux de la zone d'étude	14
2) Méthodologie	14
a) Présentation des habitats (zone humide/zone sèche) à cartographier	14
b) Carte des habitats de la zone humide	15
c) Présentation de la zone sèche.....	18
3) Réalisation d'une classification orientée objet.....	18
a) Présentation du logiciel.....	18
b) Réalisation de la classification.....	21
c) Qualité de la classification	25
4) Carte des habitats oiseaux sur l'ensemble de la zone	26
III) Premières analyses des données ornithologiques et épidémiologiques	27
1) Les données disponibles	27
a) La base de données oiseaux	27
b) Les données sérologiques	29
2) Analyse des données ornithologiques.....	30
a) Méthodologie	30
b) Cartes des abondances moyennes des oiseaux par habitat.....	31
c) Cartes du nombre d'espèces d'oiseaux par habitat	31
d) Echelle de paysage et communauté oiseaux	32
e) Impact de l'urbanisation sur la communauté oiseaux.....	33
3) Recherche de facteurs de risque West Nile.....	36
a) Analyses des variables	36
b) Analyse multi-variables	38
IV) Discussion et conclusion.....	40
a) Bilan du travail réalisé	40
b) Evolution et perspectives	40
c) Conclusion	41
Bibliographie :	42
Annexes.....	44

Introduction :

Depuis 40 ans les capacités des systèmes des satellites environnementaux ont considérablement augmenté tant en terme de résolution spatial, spectral que temporel, et l'utilisation de données de télédétection pour les études épidémiologiques a été reconnue [Ref 1]. Plusieurs travaux ont montré le potentiel d'application de la télédétection à l'épidémiologie [Ref 9, Ref 10, Ref 11]. Dans le cas d'une maladie multi-hôtes, un grand nombre d'espèces (hôtes et vecteurs) peuvent être impliquées dans le processus de transmission; la persistance et la circulation de l'agent pathogène résultant alors de l'occurrence d'associations spécifiques entre les hôtes et les vecteurs de la maladie. C'est le cas de la fièvre de West Nile dont l'émergence récente en Europe pose la question des mécanismes de transmission et les déterminants de l'amplification, la diffusion et l'émergence de la maladie, en particulier le rôle joué par la biodiversité [Ref 8].

Dans le cadre d'une étude sur la circulation du virus de West Nile en Roumanie, l'objectif du stage est de caractériser la diversité des associations dans le temps et dans l'espace des espèces d'hôtes (oiseaux sauvages), de décrire la diversité de la communauté d'oiseaux et de rechercher des facteurs de risques pour la transmission du Virus de West Nile.

Dans une première partie, nous présenterons le contexte du stage (les instituts d'accueil, le projet de recherche dans lequel s'inscrit ce travail, la problématique du virus de West Nile,). Une deuxième partie permettra d'expliquer les différentes étapes pour la réalisation d'une carte des habitats oiseaux à partir d'une carte de végétation et d'une classification supervisée d'une image Landsat. Nous expliquerons ensuite les premières analyses des données ornithologiques et épidémiologiques (partie III) et discuterons les résultats (IV).

I) Présentation

1) Le cadre d'accueil

a) Le CIRAD

Le CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement) est un Établissement Public à Caractère Industriel et Commercial (EPIC). Il est né en 1984 et est spécialisé dans la recherche agronomique appliquée aux régions chaudes. Sa mission est de contribuer au développement rural des pays tropicaux et subtropicaux par des actions de recherche, des réalisations expérimentales, des actions de formation (France et étranger) ou encore la diffusion d'information scientifique et technique. Il travaille en coopération avec plus de 90 pays d'Afrique, d'Asie, du Pacifique, d'Amérique Latine et d'Europe.

Les différents travaux sont réalisés dans des centres qui lui sont propres, et au sein de structures nationales de recherche agronomique des pays partenaires.

Le CIRAD emploie 1800 personnes dont la moitié sont des cadres de recherche en métropole ou à l'étranger. Son budget annuel, de 152 millions d'euros, provient pour les 2/3 du Budget Civil de Recherche et de Développement technologique (BCRD) et pour 1/3 de ressources contractuelles.

En France, le CIRAD possède un siège social à Paris et un centre de recherche à Montpellier, ainsi que quelques stations dans les DOM-TOM.

Les principales activités du CIRAD de Montpellier sont :

- la recherche : en interaction avec les projets du terrain. Le centre de Montpellier fournit des méthodes scientifiques (analyse des génomes, amélioration des plantes,...), des techniques (culture in vitro, biologie moléculaire, analyses minérales,...) et des outils (bases de données, statistiques,...)
- les prestations : mises à la disposition des chercheurs et des agriculteurs des régions chaudes (identification de maladies tropicales, choix de moyens de lutte, sélection variétale, analyse de bois, aliments,...)
- l'expertise : pour fournir des moyens comme la conception et l'évaluation de projets, ou encore l'appui aux politiques de recherche agronomique nationales et internationales
- la formation : le centre accueille plus de 800 chercheurs et techniciens chaque année

Il est divisé en 3 grands départements de recherche : « Systèmes biologiques », « Performances des systèmes de production et de transformation tropicaux » et « Environnement et société ». Chaque département est composé de plusieurs unités de recherche. Le département systèmes biologiques comporte 23 unités de recherches dont l'UPR16 : « Epidémiologie et écologie des maladies animales » à laquelle j'ai été rattaché.

b) L'unité Epidémiologie et écologie des maladies animales

Mon stage s'est déroulé au sein de l'unité « Epidémiologie et écologie des maladies animales » (UPR16). Cette unité est essentiellement basée sur le site Baillarguet à Montpellier, et est formée d'épidémiologistes, entomologistes, écologues, de vétérinaires et de spécialistes en télédétection et Systèmes d'Information Géographique. L'objectif des recherches est l'évaluation de l'importance et de l'impact des maladies animales et de l'anticipation de l'émergence de celles-ci afin de décider des méthodes de prévention et de lutte. Un deuxième objectif est l'étude des mécanismes de transmission de ces maladies et en

particulier l'identification des déterminants environnementaux favorables à leur développement.

L'axe des recherches est essentiellement tourné vers la trypanosomose animale, la bluetongue (fièvre catarrhale ovine), la fièvre de la vallée du Rift, la grippe aviaire, la fièvre porcine et le virus de West Nile.

Les principales activités de l'unité porte sur la mise en œuvre d'étude d'observation en écologie et épidémiologie qui implique un aspect entomologie (bio-écologie des vecteurs, génétique), géomatique (télédétection, SIG, statistiques spatiales), modélisation statistique et mathématique.

c) La Maison de la Télédétection

La Maison de la Télédétection est une structure régionale qui est entrée en fonction en septembre 1994. Elle regroupe des unités de recherche de différents instituts pour constituer un pôle de recherches appliquées en télédétection et information géographique. La Maison de la Télédétection m'a accueilli pendant toute la durée de mon stage.

La principale unité de recherche basée à la maison de la télédétection est l'UMR TETIS « Territoires, Environnement, Télédétection et Information Spatiale ». Elle rassemble 70 agents permanents du CIRAD, du Cemagref et de l'ENGREF ainsi qu'une vingtaine de doctorants et une dizaine de personnes sur contrats. Trois membres de l'UPR16 Epidémiologie et écologie des maladies animales y sont détachés.

Son objectif général est de produire, développer, diffuser et transférer des connaissances, des théories et des méthodes permettant de décrire et comprendre les structures et dynamiques spatio-temporelles des écosystèmes, des espaces ruraux et des territoires, et également de maîtriser l'information spatiale au service des démarches de connaissance, de gestion agri-environnementale et de développement territorial.

d) Le projet EDEN

Le projet EDEN (Emerging Diseases in a changing European eNvironment) est un projet intégré européen, dont l'objectif est d'identifier les écosystèmes et les conditions environnementales qui influencent la dynamique spatiale et temporelle d'agents pathogènes humains, ainsi que de développer et coordonner un ensemble de méthodes génériques, et d'outils d'aide à la décision pour les politiques de santé publique.

EDEN regroupe des chercheurs de 48 institutions, réparties dans 24 pays (Cf. Figure 1). Le projet couvre ainsi l'ensemble des écosystèmes européens, depuis le cercle polaire au Nord jusqu'au bassin Méditerranéen (avec des liens en Afrique de l'ouest) au Sud, depuis le Portugal à l'Ouest jusqu'au delta du Danube à l'Est.

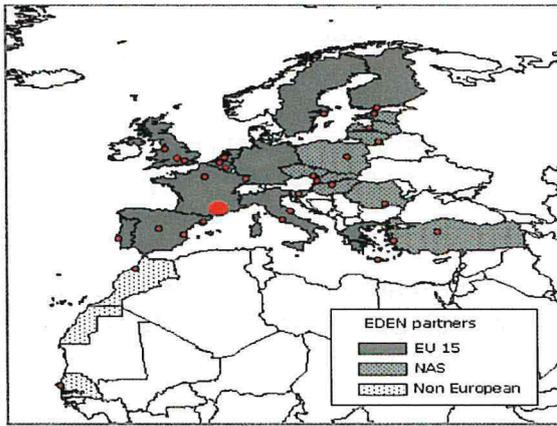


Figure 1 : Répartition des institutions participant au projet EDEN

Différentes maladies humaines, particulièrement sensibles aux changements environnementaux, sont étudiées dans un cadre scientifique commun. Les recherches sont organisées suivant 6 axes : le virus du Nil Occidental, les maladies à tiques, les virus liées aux rongeurs, la malaria, les leishmanioses et les maladies africaines.

Le projet EDEN, coordonné par le Cirad, intègre les activités de recherche des différents partenaires, par une organisation en sous projets par maladie liés par un ensemble d'activités d'intégration comprenant l'étude de la biodiversité, la détection des changements

environnementaux, la modélisation épidémiologique, le traitement d'images de télédétection, l'information et la communication.

Dans le cadre du stage, j'ai été intégré au projet EDEN au sein de l'axe West Nile pour apporter ma collaboration dans le domaine de la télédétection et des Systèmes d'Information Géographique pour l'étude du virus de West Nile en Roumanie.

2) Le virus de West Nile

a) Historique

Le virus de West Nile ou virus du Nil Occidental en français a été isolé pour la première fois en 1937 en Ouganda, dans le district de West Nile (d'où il tire son nom) [Ref 7]. Il a ensuite été détecté chez des hommes, des oiseaux et des moustiques en Egypte au début des années 50 [Ref 7], et a depuis été recensé chez l'homme ou l'animal dans divers pays. Des cas humains de fièvre de virus de West Nile ont été rapportés en Afrique, au Moyen Orient, en Inde, en Europe et plus récemment sur le continent américain, où une première épidémie s'est déclarée dans la ville de New York en 1999 [Ref 5]. Des épidémies limitées ont eu lieu aux Etats-Unis les deux années suivantes, et c'est en 2002 que le virus s'est considérablement répandu, touchant 4156 personnes dans 44 états et faisant 284 décès. En 2003, les Etats-Unis recensaient plus de 9000 cas et 264 décès. C'est la plus grande épidémie de fièvre à virus West Nile jamais répertoriée. L'année 2003 avait également été marquée par l'apparition d'infections par le virus West Nile au Mexique (591 cas au 10 décembre 2003). Un léger fléchissement marque l'année 2004, avec 2470 infections humaines et 88 décès aux Etats-Unis.

En France, des cas d'infection sont apparus chez des chevaux en Camargue en 2000, et 7 cas humains ont été recensés en 2003 dans le Var. Toujours en Europe, une importante épidémie a eu lieu en Roumanie en 1996-97, à Bucarest où plus de 500 personnes furent infectées avec une mortalité de 10%. En juillet 2004, deux cas humains étaient rapportés dans le sud du Portugal.

Au cours des 10 dernières années le Virus de West Nile a montré sa capacité à provoquer des épidémies de plusieurs centaines de méningo-encéphalites dans le bassin méditerranéen, mais surtout aux Etats-Unis. Cependant, comparé à d'autres flavivirus, comme la dengue, son importance sanitaire reste relative. Mais il est sans doute l'arbovirus le plus largement

distribué, manifestant de grandes capacités d'adaptation à de nombreux écosystèmes [Ref 6, Ref 7].

b) Système épidémiologique

Le virus de West Nile est un virus de la famille des Flaviviridae (flavivirus), porté par un vecteur arthropode suceur de sang (arbovirus), il s'agit généralement d'un moustique du genre *Culex* (*C.pipiens* ou *C.Modestus* en Europe) [Ref 2]. Le cycle de transmission du virus est formé par les oiseaux sauvages comme hôtes principaux et les moustiques comme vecteurs [Ref 2]. Les oiseaux constituent le réservoir épidémiologique et les moustiques, le vecteur amplificateur. Comme l'illustre le schéma de la Figure 2, un vecteur au départ sain va devenir positif en piquant un hôte contaminé. Ce moustique va amplifier la maladie et sera susceptible de la retransmettre à un autre hôte. Ce dernier sera à son tour porteur de la maladie. Certaines espèces d'oiseaux ne présentent aucun signe apparent de contamination mais d'autres peuvent en mourir.

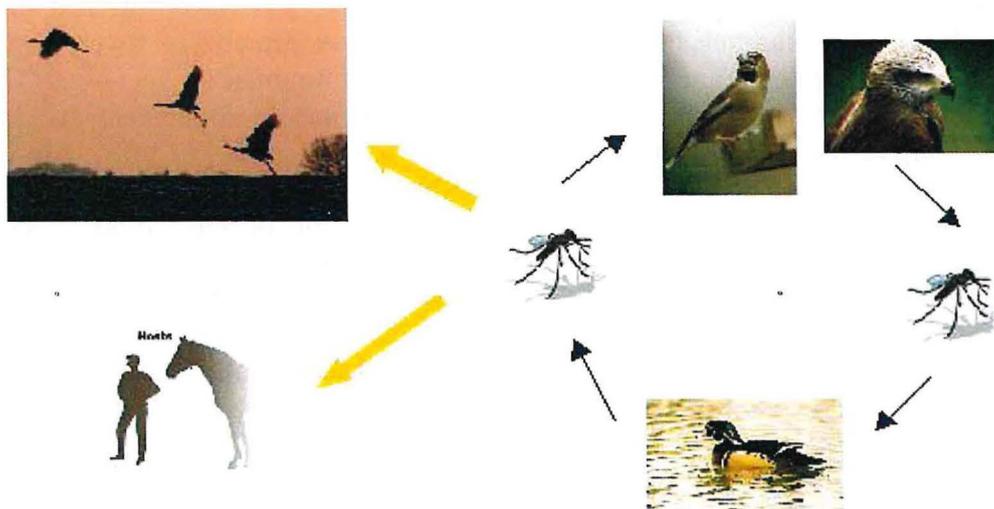


Figure 2 : Système épidémiologique du virus de West Nile (le choix des espèces d'oiseaux représentées n'a pas de signification épidémiologique particulière)

Les mammifères, dont l'homme et le cheval sont des hôtes accidentelles, considérés comme des hôtes cul de sac [Ref 2]. Le cheval et l'homme ne font pas partie du cycle principal du virus, ils font l'objet d'une transmission accidentelle et sont considérés comme des « culs de sac épidémiologiques ». En effet, leur virémie est courte et de faible amplitude, ce qui ne permet pas d'infecter des moustiques vecteurs potentiels [Ref 7].

L'homme et les équidés sont sensibles à l'infection avec une majorité de cas non symptomatiques. Le virus peut cependant dans certains cas provoquer des infections aiguës du système nerveux central notamment des épidémies de méningo-encéphalite parfois mortelle chez l'homme ou des épizooties chez les chevaux.

Les chevaux infectés peuvent manifester des signes de fièvre suivis d'encéphalomyélites et vacillements, paralysie du train arrière amenant souvent à la mort [Ref 2].

3) Contexte du stage

Le cycle du virus West Nile s'avère extrêmement complexe, faisant potentiellement intervenir une multitude d'espèces de moustiques et un grand nombre d'espèces d'oiseaux différentes. Parmi ces oiseaux, on note des oiseaux migrateurs qui ont pu être infectés dans un pays, et retransmettre la maladie dans un autre pays.

Etant donné que le réservoir de la maladie est très mobile, l'étude du virus sur une petite zone a peu de sens. Un des objectifs du projet EDEN (pour le virus de West Nile) est d'élargir la zone de surveillance à l'échelle de l'Europe et à une partie du Sénégal pour pouvoir suivre l'évolution géographique du virus. L'idée est donc d'uniformiser les méthodes de travail de chaque institut de recherche concerné et ainsi avoir des données comparables dans chacune des zones d'études. La problématique actuelle est de déterminer s'il existe un écosystème favorable au développement du virus. On sait que les principaux foyers des infections du virus de West Nile sont principalement situés dans les écosystèmes humides (deltas, zones inondables) [Ref 2]. Etant donné que le système épidémique est principalement basé sur les moustiques et les oiseaux, il paraît évident que le milieu le plus favorable soit un milieu humide et chaud. On remarque effectivement que les deux principaux foyers du virus en Europe sont la Camargue en France et le delta du Danube en Roumanie.

Pour essayer de connaître les habitats les plus favorables au développement du virus, une étude a été réalisée sur la Camargue, avant mon arrivée en stage. En effet, une carte de l'occupation du sol représentant les principaux habitats de la communauté oiseaux et moustiques a été réalisée. Une base de données décrivant la répartition, l'activité et l'abondance des espèces d'oiseaux présentes en Camargue et potentiellement impliquées dans la circulation du virus de West Nile a été réalisée par l'équipe ornithologique française partenaire du projet EDEN.

L'ensemble de ces informations réunies sur un SIG s'avère être un outil précieux pour la compréhension du cycle du virus. Il permet notamment de connaître les principaux points de contact (moustiques/oiseaux) en réalisant des cartes d'abondance pour des espèces données. Ces données peuvent être couplées aux données sérologiques sur le cheval (ou sur l'homme) pour essayer de déterminer les espèces de moustiques, ou le type d'habitat favorable au développement du virus. Les outils de géomatique apportent ainsi une aide capitale pour le suivi et la compréhension du développement de la maladie.

Maintenant que cet outil est disponible en Camargue, un des objectifs d'EDEN est d'uniformiser le travail réalisé dans les autres centres partenaires ce qui permettra d'avoir des données comparables sur l'ensemble de la zone d'étude. Le delta du Danube est le principal foyer du virus en Europe, il est donc important de développer cet outil pour ce site. Etant donné que les membres du Danube Delta National Institute partenaires au projet EDEN n'étaient pas en mesure de réaliser ce travail, il a été décidé d'effectuer un partenariat avec l'équipe roumaine pour réaliser ce projet. C'est cette mission qui a fait l'objet de mon stage dont l'objectif est de réaliser avec l'équipe ornithologique une carte des habitats oiseaux, permettant ainsi de caractériser la diversité et la distribution de la communauté oiseaux. Une étude épidémiologique pourra ensuite être réalisée en couplant les données obtenues avec des données sérologiques sur la circulation du virus de West Nile.

4) La zone d'étude

La zone d'étude est située à l'Est de la Roumanie, elle comprend une grande partie de la région de Dobrogea, de Brăila jusqu'à la mer Noire et de la frontière ukrainienne jusqu'à la pointe sud du delta du Danube (à 30 km au nord de Constanta) (Cf. *Figure 3*).

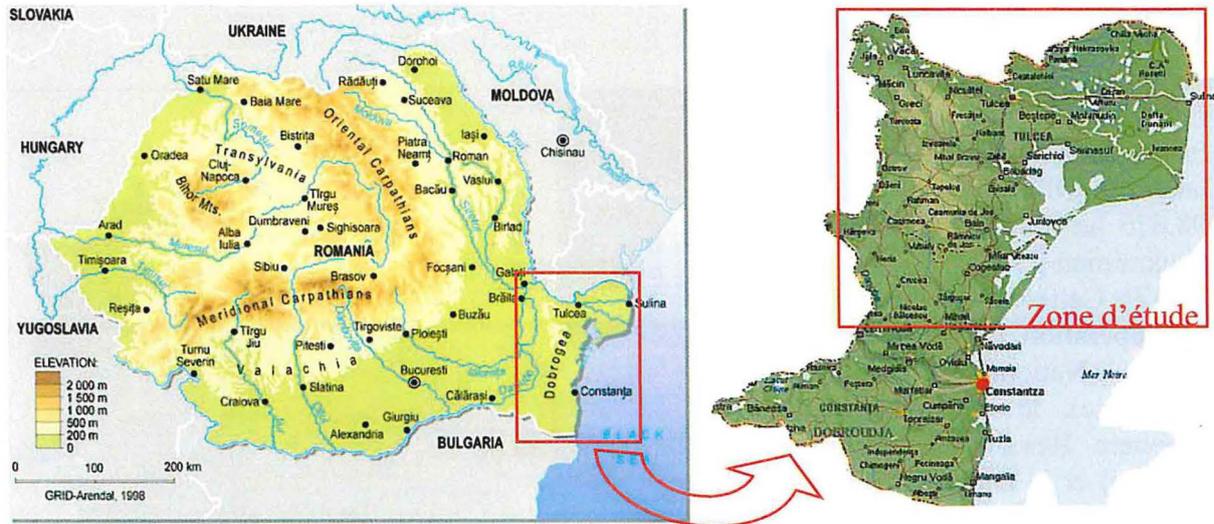


Figure 3 : Localisation de la zone d'étude

S'étendant sur plus de 140 km d'Est en Ouest, notre zone d'étude comprend une zone sèche : le plateau de Dobrogea et une partie humide : le delta du Danube (partie roumaine). Ce delta est le plus grand et le mieux préservé delta d'Europe avec 2733 km² classés espace protégé. Cette réserve naturelle recense 300 espèces d'oiseaux différentes, dont de nombreuses espèces migratrices.

Le plateau de Dobrogea est à l'inverse une zone extrêmement sèche, avec le taux de pluviométrie le plus faible du pays. L'activité principale est l'agriculture avec des cultures céréalières en forte majorité. Bien que le plateau de Dobrogea soit une région beaucoup plus sèche avec une faune un peu moins riche, il n'en reste pas moins une zone essentielle à prendre en compte pour l'étude car il recense la plupart des villages et des élevages de chevaux étant potentiellement exposés au virus.

II) Réalisation d'une carte des habitats oiseaux

1) Les données utilisées

a) Carte de Végétation

On dispose d'une carte de végétation de l'ensemble de la réserve de biosphère du delta du Danube (DDBR) qui couvre 398.676 ha dont 35.711ha sur le sol ukrainien [

Ref 12]. Cette carte a été réalisée en coopération par le « Danube Delta National Institute » (Roumanie), le « Danube Delta Biosphere Reserve Authority » (Ukraine) et le Dutch Directorate General for Public Works en 2002. La carte consiste en une mosaïque de différentes unités de végétation avec une échelle approximative des polygones de 1 : 140.000. Elle est composée de 46 unités de végétation. L'ensemble de ces classes est organisé dans 9 grands ensembles : végétation côtière (plage, dunes...), steppes et dunes (non côtières), steppes désertiques, prairies en bordures de rivières, végétations marécageuses, forêts et buissons sur dunes, forêts et buisson sur terrains inondables, végétations aquatiques et une classe divers (polders agricoles, étang à poisson (élevage) et zones urbaines) (Cf. Figure 4).

La réalisation de cette carte a fait l'objet d'un rapport dans lequel est décrit les différentes unités de végétation de la carte avec le nom scientifique des différentes espèces dominantes de chaque habitat [Ref 12].

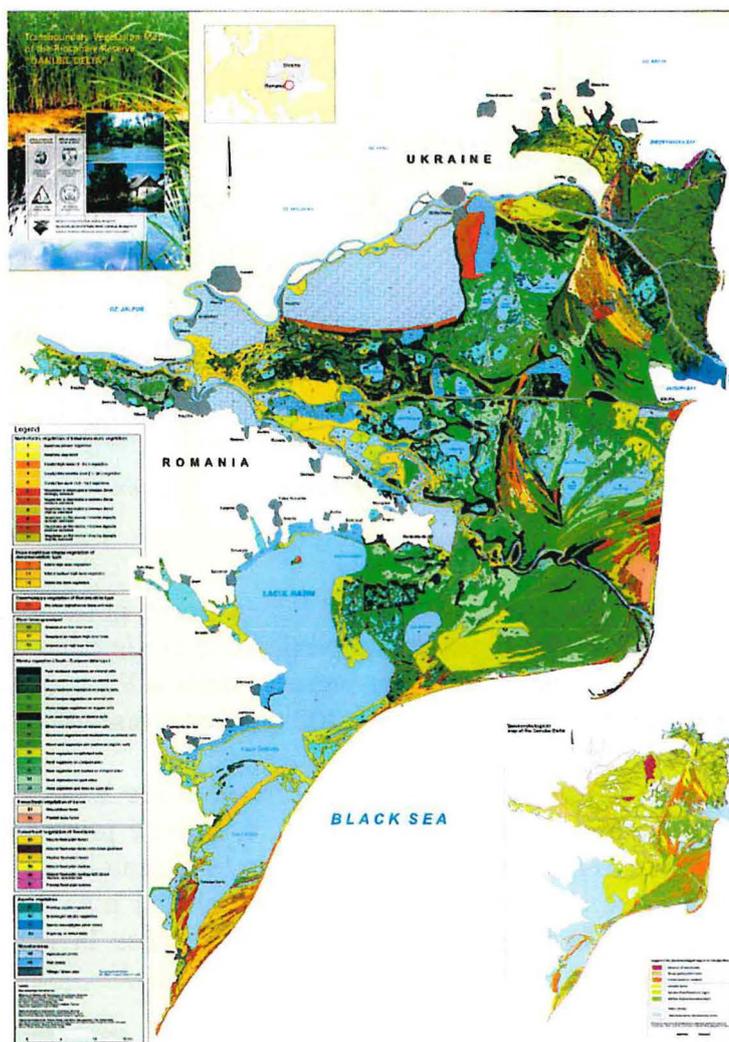


Figure 4 : Carte de végétation de la réserve de la biosphère du delta du Danube (DDBR)

b) Image Landsat

On dispose d'une image prise par le satellite Landsat 5, le 31 mai 2006. Celle-ci est composée de 7 canaux : TM1 à TM3 dans le domaine du visible (bleu, vert, rouge), TM4 dans le proche infrarouge, TM5 et TM7 dans l'infrarouge moyen et TM6 dans l'infrarouge thermique.

L'image englobe la quasi-totalité de notre zone d'étude, seule une partie au nord-est est manquante. Cette absence de données n'a pas d'importance, car étant donné que l'on dispose d'une carte de végétation (présentée précédemment) aucun travail de télédétection n'est envisagé sur cette zone. L'image a volontairement été choisie au printemps qui correspond à la période où l'activité chlorophyllienne des plantes est la plus forte (climat continental), permettant ainsi de dissocier plus facilement différents couverts végétaux. La résolution spatiale de l'image est de 30 mètres au sol (pas de mode panchromatique disponible sur Landsat 5).

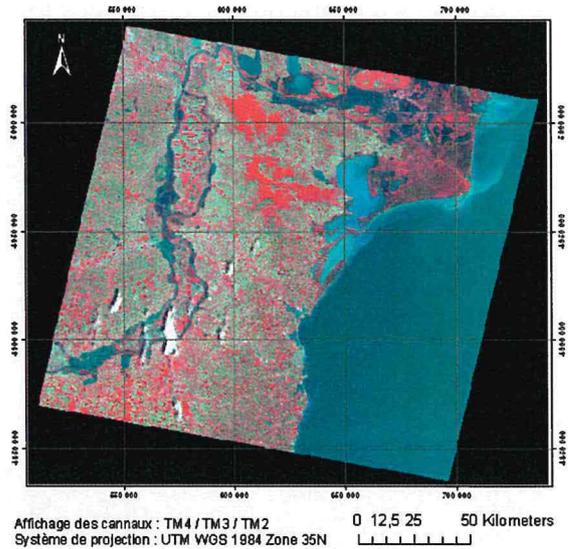


Figure 5 : Extrait d'une scène du satellite Landsat 5, composition colorée fausse couleurs (TM4-R, TM3-V, TM2-B).

c) Modèle Numérique de Terrain

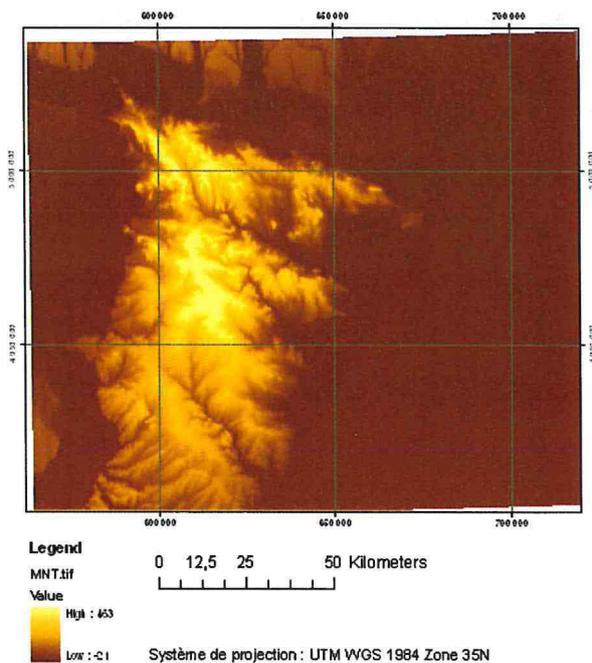


Figure 6 : MNT SRTM d'une résolution de 90 mètre redimensionné sur la zone d'intérêt

Nous utiliserons également un Modèle Numérique de Terrain (MNT) de la Shuttle Radar Topography Mission d'une résolution de 90 mètres. Ces données présentent l'avantage d'être mises à disposition gratuitement sur le site de la NASA et leur résolution est tout à fait suffisante pour notre application. Notre zone d'étude étant à cheval sur deux images SRTM, nous procéderons à une fusion des deux images puis à un découpage afin de ne garder la partie qui nous intéresse (Cf. *Figure 6*).

d) Données collectées sur le terrain

Les données terrains ont été collectées lors de ma mission réalisée du 9 au 24 juin 2007. Cette mission a été composée d'une partie terrain dans le delta du Danube, d'une partie terrain en zone sèche et d'une partie au Danube Delta National Institute.



Figure 7 : Ensemble des relevés terrains effectués, localisés par GPS (points numérotés en jaunes). Le fond de carte utilisée est l'extrait de la scène Landsat du 31 mai 2006 (composition colorée fausse couleur TM4-TM3-TM2)

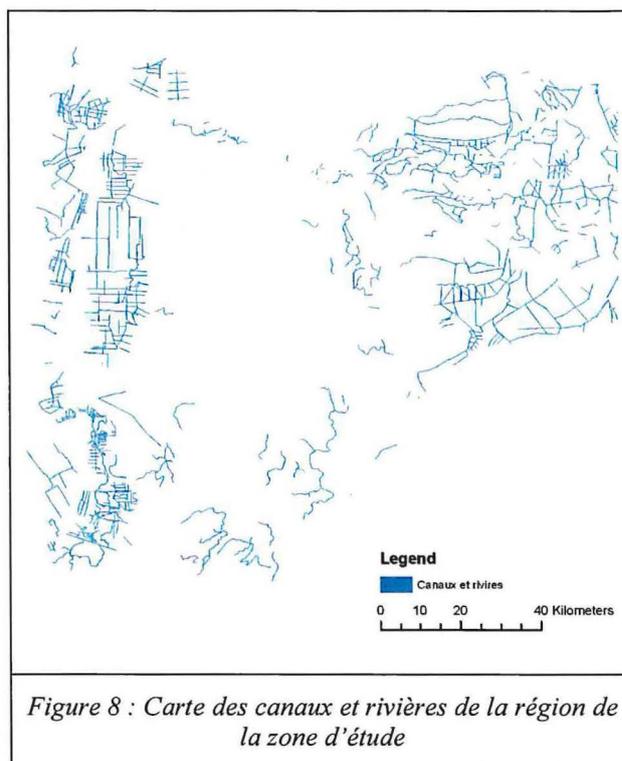
La partie terrain en zone humide a été principalement réalisée en compagnie de Mihai Marinov, un ornithologue. Cette mission en durée environ 6 jours (en trois fois) au moyen d'un bateau. Cette mission m'a permis de me familiariser avec les différents habitats de la zone humide et de prendre de nombreux relevés terrains, localisés par GPS (Cf. Figure 7). Etant donné que l'on dispose d'une carte de végétation sur cette zone, ces relevés ont été utilisés pour effectuer des vérifications et illustrer la liste des habitats (Cf. Annexe 1).

La partie terrain en zone sèche a été réalisée en compagnie de Mihai Doroftei, un botaniste. La visite des différents habitats s'est effectuée en voiture et à pieds pour les zones les moins accessibles. Pour chaque habitat de nombreux relevés terrains ont été effectués. Chaque relevé est composé de 2 points GPS au minimum permettant de délimiter une zone couvrant plus de 3*3 pixels, d'un croquis explicatif et des espèces végétales dominantes (Cf. Fiche terrain Annexe 3).

Ces relevés de végétation de la zone sèche ont été utilisés pour réaliser une classification dirigée à partir de l'image Landsat (présentée précédemment).

e) Carte des rivières et canaux de la zone d'étude

On dispose d'une carte de l'ensemble des canaux et rivières de la région de Dobrogea (Cf. Figure 8). Cette carte provient du Danube Delta National Institute, et est disponible au format shape (format de vecteurs pour ESRI ArcGis). Cette carte est importante car les canaux attirent un grand nombre d'espèces d'oiseaux, mais la plupart d'entre eux sont trop étroits pour être détectés sur une image satellite de résolution moyenne comme Landsat.



2) Méthodologie

a) Présentation des habitats (zone humide/zone sèche) à cartographier

La liste des habitats représentatifs pour décrire la communauté oiseaux a été principalement effectuée par l'équipe ornithologique du Danube Delta National Institute, en collaboration avec un botaniste. Ces habitats n'ont pas comme objectif de décrire l'écosystème de la zone étudiée mais d'énumérer les principales zones qui décrivent au mieux la répartition de la communauté oiseaux en fonction de leur activité. Cette liste est susceptible d'être modifiée car créer une base de données oiseaux sur une zone aussi grande est difficile et demande énormément de temps sur le terrain.

Au total, 32 habitats différents sont utilisés pour décrire la répartition des oiseaux sur l'ensemble de la zone d'étude (Cf. Tableau 1). Parmi ces habitats, on peut différencier :

- **Les habitats propres à la zone humide :** zones agricoles (delta), zones urbaines et rurales, eau côtière, lagunes, zones côtières, marécages d'eau douce, marécages, forêts (chênes sur sol marin), forêt naturelle inondée, forêt générale, buissons, prairies, steppes, roselières avec quelques formations de saules et clairs de roselières.
- **Les habitats de la zone sèche :** zones agricoles sur le plateau, verger, vignes, zones rocheuses, carrières, zones urbaines et rurales, petites rivières en zones sèche, forêt en zone sèche, buisson en zone sèche.
- **Les habitats communs aux 2 zones :** étangs à poissons (élevage), eau courante, eau claire, ripisylves, roselières.

Tableau 1 : Liste des habitats oiseaux (zone humide et sèche)

No	Description	traduction
1	Agricultural areas (polders)	Zones agricoles (polders)
2	Agricultural Area on Tableland (cereals)	Zones agricoles en zone sèche (céréales)
3	Orchard on Tableland	Verger en zone sèche
4	Vineyard	Vignes
5	Fishponds	Etang à poissons (élevage)
6	Rocky Area on Tableland	Zones Rocheuses
7	Quarry	Carrières
8	Urban and rural settlement on Tableland	Zones urbaines et Rurales à l'intérieur du delta
9	Urban and rural settlement in DDBR	Zones urbaines et Rurales en zone sèche
10	Running water	Eau courante
11	Small river on Tableland	Petite rivière en zone sèche
12	Clear Water	Eau claire
13	Coastal marine water	Zone côtière (coté mer)
14	Lagoon	Lagunes
15	Coastal marine zone	Bords de mer (coté terre)
16	Fresh water marshes	Marécage à eau douce
17	Marshes	Marécages
18	Forest (mixed oak wood on marine levee)	Forêt (chênes sur sol marin)
19	Natural floodplain forest	Forêt naturelle inondable
20	Riparian forest	Ripisylves
21	Forest on tableland	Forêt en zone sèche
22	Forest (other)	Forêt (autre)
23	Bushes	Buissons
24	Bushes on tableland	Buissons en zone sèche
25	Tableland's foot	Talus
26	Marine levee	digue marine
27	Meadows	Prairie
28	Steppes in DDBR	Steppes à l'intérieur du delta
29	Steppe on Tableland	Steppes en zone sèche
30	Reed beds (with small willow formations)	Roselières avec quelques formations de saules
31	Reed beds (other)	Roselières (autre)
32	Floating aquatic vegetation	Clairs de roselières

Il s'est avéré utile de dissocier pour certains habitats comme « zones urbaines » ceux qui faisaient partie de la zone humide et ceux qui faisaient partie de la zone sèche, car ils sont inclus dans des écosystèmes différents. Les espèces d'oiseaux associées à ces classes sont donc différentes.

b) Carte des habitats de la zone humide

Pour la réalisation de la carte des habitats oiseaux, nous disposons d'une carte de végétation sur l'ensemble du delta (Cf. II) 1) a)). Cette carte est relativement précise et décrit relativement bien la plupart des unités écologiques de la liste des habitats. De nombreuses

classes de la carte des habitats ont été directement créées par association de classes de la carte de végétation (Cf. la liste des associations, Annexe 2).

Certaines classes, cependant, s'avèrent importantes pour la communauté oiseaux, mais ne constituent pas une unité de végétation spécifique. Il n'est donc pas possible de les obtenir directement avec la carte de végétation. Ces classes sont les suivantes : canaux, ripisylves, bord de mer, zone côtière, clairs de roselières, marécages, marécages à eau douce et également les talus (tableland's foot) dont une partie est située à la limite de la zone sèche et du delta. Ces classes ont du être créées pour pouvoir être ajoutées à la carte des habitats. Les méthodes utilisées pour représenter ces classes sont explicitées ci-dessous.

Le travail d'association et de création de classe a été effectué à l'aide du SIG ArcGis (ESRI). Ce logiciel présente l'avantage de pouvoir manipuler facilement des objets vecteurs au format shape. Les principales fonctions utilisées dans cette partie appartiennent à la boîte d'outils « Analysis Tools », il s'agit de fonctions de découpages, de superpositions et de calculs des zones tampons.

La cartographie des canaux n'est pas envisageable par télédétection avec une image de résolution moyenne (Landsat) car certains sont trop petits pour être visible sur l'image. Il a donc été décidé d'utiliser la carte la carte des canaux et rivières présentée précédemment. Cette carte sera utilisée pour le delta où les canaux sont associées à l'eau courante (running water) et pour la zone sèche où ils seront nommés petites rivières (small rivers on tableland).

Les ripisylves correspondent aux forêts le long des cours d'eau. Elles sont en général relativement étroites (30 mètres) et donc non détectable avec des images de résolution moyenne. Cet habitat joue cependant un rôle très important pour les oiseaux et il est impératif de le cartographier. Etant donné que ces forêts sont disposées sur le bord de la plupart des cours d'eau, il a été décidé de créer une zone tampon de 30 mètres de large autour de ceux-ci pour représenter cette classe.

Il en est de même pour les classes « bords de mer » et « zones côtières » où un buffer de 250 mètres a été appliqué sur l'intérieur des terres pour le premier et un buffer de 5 km sur la mer pour le deuxième (au delà de cette zone, on considère qu'il n'y a plus d'oiseaux en concentration suffisante pour être représentés).

Les clairs de roselières correspondent à une étendue d'eau libre à l'intérieur des roselières. Cette classe a été déterminée par télédétection en détectant les zones moins végétalisées autour des roselières. Cette étape a été effectuée sur eCognition au moment de la classification (Cf. II) 3) c) Réalisation de la classification).

Les talus (tableland's foot) sont situés tout autour du plateau, il s'agit d'une forte pente sur une hauteur de un à cinq mètres. Ils sont disposés à une altitude faible (inférieur à 15 mètres) ce qui permet de les différencier facilement des collines ou des zones rocheuses. Cette classe a été déterminée par seuillage à l'aide du MNT SRTM et des pentes calculés à partir de celui-ci. Bien entendu, les décrochements ont une largeur très inférieure à la résolution du MNT, mais ce qui nous intéresse ici c'est de connaître la position approximative des oiseaux concernés par cet habitat, une résolution de 90 mètres est alors suffisante.

La superposition de ces classes et de celle obtenues à partir de la carte de végétation permet de dessiner la carte des habitats oiseaux de la zone humide (Cf. *Figure 9*).

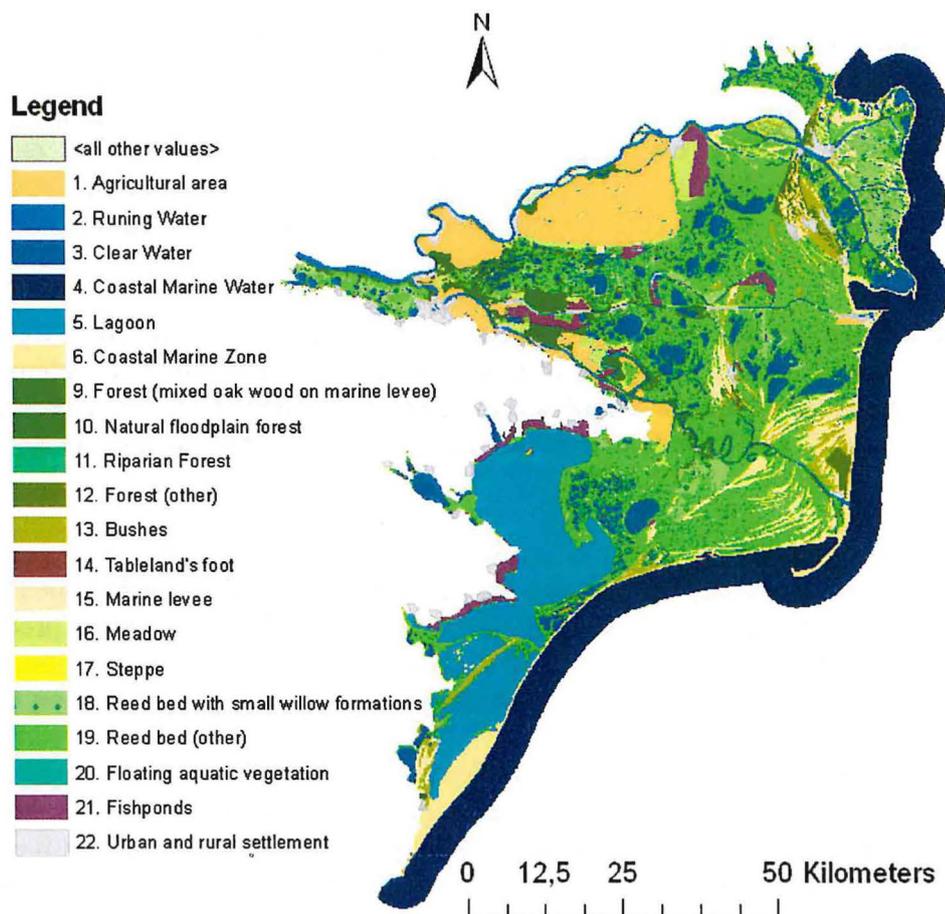


Figure 9 : Carte des habitats oiseaux pour le delta du Danube

Remarque :

Une des difficultés rencontrées n'a pas pu être résolue pendant la durée de mon stage, il s'agit des classes des zones marécageuses. En effet ces classes n'apparaissent pas dans la carte des habitats car il n'a pas été possible de les localiser précisément. Les zones marécageuses posent un problème pour la cartographie car le niveau de l'eau dans le delta du Danube varie facilement de 2 à 3 mètres, ce qui signifie que suivant les années (même à la même saison) ces zones sont parfois inondées et parfois sèches durant une année entière. Durant une période sèche, les marécages se laissent facilement envahir par la végétation (herbes, buissons...) et sont parfois même utilisés comme terres agricoles. Les populations d'oiseaux sont bien entendu très sensibles à ces changements, ce qui signifie que la même zone peut être considéré comme 4 ou 5 habitats différents suivant le niveau des eau, donc suivant la saison ou l'année. Il a finalement été décidé de créer une zone de marécages qui correspond aux zones régulièrement inondées (au moins une fois tout les deux ans). La cartographie de cette zone doit être effectuée avec par télédétection à l'aide de deux images satellites : l'une correspondant à une période sèche et l'autre correspondant à période où le niveau est relativement haut. Ce travail n'a pas pu être réalisé car je ne disposais que d'une seule image satellite.

c) Présentation de la zone sèche

La zone sèche comprend une grande partie du plateau de Dobrogea. Pour cette zone, aucune carte de végétation n'est disponible, on utilisera les outils de télédétection pour réaliser une cartographie de l'occupation du sol. Les relevés de végétation qui ont été réalisés lors de la mission de terrain, serviront à guider la classification.

Certaines classes, comme les rivières et les talus, ne peuvent pas être détectées par télédétection avec une image de résolution spatiale moyenne. Ces classes seront déterminées à partir de la carte du réseau hydrologique et du Modèle Numérique de terrain, de la même manière que pour la zone humide (Cf. II) 2) b)). Ces classes seront ensuite superposées à la classification finale à l'aide du SIG ArcGis.

3) Réalisation d'une classification orientée objet

Pour réaliser la carte des habitats oiseaux de la zone sèche, nous avons réalisé une classification supervisée à l'aide du logiciel eCognition. Ce logiciel permet d'effectuer des classifications orientées objet, présentant des fonctionnalités particulières. Dans un premier temps, nous présenterons le logiciel, nous expliquerons ensuite les différentes étapes de la classification, puis nous évaluerons sa qualité à l'aide d'une matrice de confusion.

a) Présentation du logiciel

Le logiciel eCognition est un logiciel d'analyse d'image, développé par Definiens, une société privée. A l'origine ce logiciel était adapté à l'imagerie médicale puis s'est généralisé à l'étude d'images de télédétection.

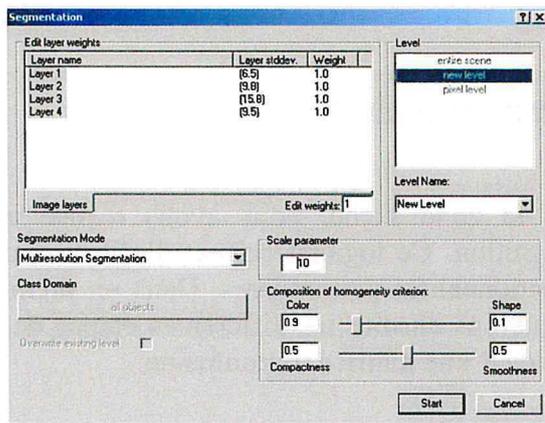
Sa spécificité est la description des surfaces par une approche orientée objet. C'est-à-dire que l'utilisateur travaille sur des objets radiométriquement « homogènes » et non pas sur les pixels eux-mêmes. Cette approche permet à l'utilisateur d'avoir une multitude de paramètres supplémentaires pour guider sa classification. Par exemple, il peut tenir compte des paramètres d'homogénéités, de forme des objets (élongation, aire, forme circulaire...), ainsi que des paramètres de proximité aux autres classes.

i) Segmentation

Afin de passer d'une information raster à une information vectorielle, il est nécessaire de segmenter l'image. Cette première étape est une particularité de la classification orientée objet. La segmentation multi résolution permet de générer des objets de tailles variables (structures fines ou grossières) selon le choix de l'utilisateur.

Deux modes de segmentation sont disponibles :

- La procédure de segmentation dite Normale (Multiresolution Segmentation) est basée sur une technique ascendante de croissance de région commençant par un objet constitué d'un seul pixel de l'image. Durant le processus les objets sont fusionnés par paires pour créer des objets de taille supérieure tout en réduisant au minimum l'hétérogénéité. A chaque itération sont fusionnés les deux objets adjacents de l'image ayant une fois réunit la plus faible croissance d'hétérogénéité. Si la plus faible croissance d'hétérogénéité excède le seuil défini par le paramètre d'échelle, le processus s'arrête.
- Le mode de segmentation « Différence Spectrale » fonctionne selon le même procédé en prenant en compte uniquement des valeurs spectrales de l'image



L'utilisateur joue sur des paramètres d'échelle (Cf. **Figure 10**), de forme et de compacité pour créer les éléments porteurs de l'information.

Figure 10 : Interface de segmentation sous eCognition

Le Paramètre de Forme / Couleur et de Compacité / Lissé font partie des critères d'Homogénéité. Le critère de couleur définit le pourcentage de l'homogénéité globale issu des valeurs spectrales des canaux de l'image par opposition aux critères d'homogénéité de forme. Ce dernier est à employer dans le cas d'image radar (données très texturées) afin d'éviter une forme trop « fractale » des objets. En plus de l'information spectrale, eCognition permet d'optimiser selon la forme des objets a bordures compactes (rapport entre la longueur de la bordure et la racine carrée du nombre de pixels formant l'objet) ou lisses (rapport entre sa longueur et la longueur du plus petit côté).

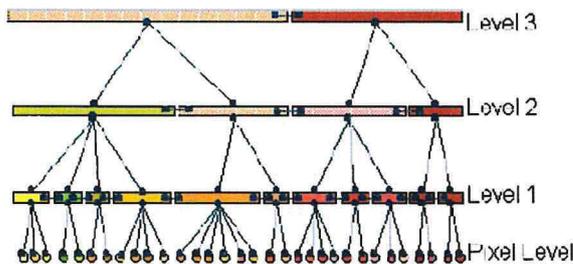


Figure 11 : Schéma de l'imbrication de différents niveaux

Il est possible de pondérer l'intervention des couches en entrée. Le système de Niveau permet de créer des segmentations imbriquées (Cf. Figure 11) de plus grande ou plus petite taille. eCognition permet aussi d'intégrer des couches thématiques pour affiner la segmentation.

ii) Classification

eCognition propose deux techniques de classification supervisée : l'approche par plus proche voisin et les fonctions d'appartenance. Les deux approches peuvent être combinées pour constituer la stratégie la plus appropriée pour chaque classe.

La classification est conduite en utilisant la logique floue. La classification floue assigne non seulement un objet d'image à une classe mais elle fournit en plus la probabilité d'affectation à toutes les classes considérées.

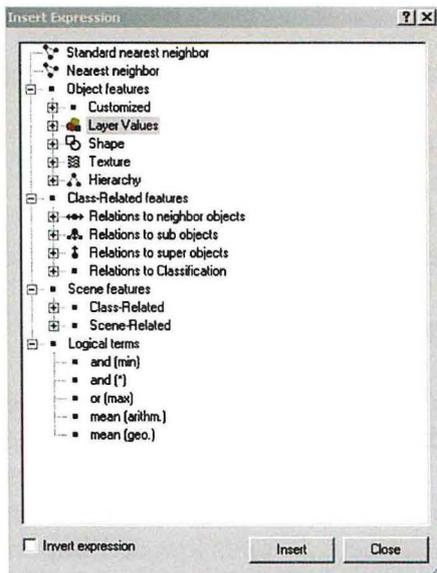


Figure 12 : Les attributs eCognition

Les fonctions d'appartenance (Cf. Figure 13) constituent une méthode simple pour traduire une gamme arbitraire d'attributs en une gamme uniforme comprise entre 0 et 1. Une valeur d'appartenance entre 0 et 1 est renvoyée par la fonction d'appartenance après évaluation de chacun des attributs formant une description de classe. Ces valeurs sont combinées en une valeur globale d'affectation de classes au moyen d'opérateurs logiques.

L'édition graphique est limitée aux fonctions d'appartenance à une dimension, leur utilisation est recommandée lorsque une classe peut être séparée des autres par un nombre limité de critère.

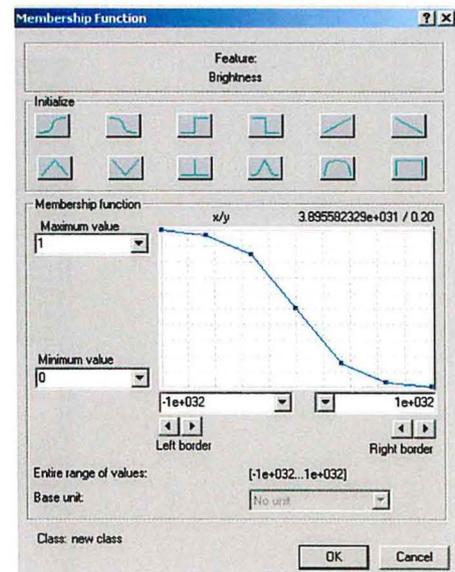


Figure 13 : Fonction d'appartenance eCognition

La classification orientée objet permet la création de différents niveaux qui sont donc liés les uns aux autres, on peut ainsi définir des relations de classification entre objets de même niveau mais aussi entre objets de niveau inférieur ou supérieur.

Le langage orienté objet permet aussi de différencier hiérarchie d'héritage et de groupes. La première notion concerne la transmission des descriptions de classes vers des classes filles et la deuxième de regrouper sémantiquement des classes aux descriptions d'attributs variés.

b) Réalisation de la classification

La classification sous eCognition s'effectue en plusieurs étapes. En premier lieu, il faut créer un projet en incluant l'ensemble des couches désirées, puis commencer la classification en procédant par étape. A chaque étape, il est important de voir quelles sont les objets qui sont bien classés et ceux qui le sont moins afin d'apporter les modifications nécessaires.

i) Création du projet

Durant ma mission terrain, je me suis rendu compte que le relief avait énormément d'impact sur l'occupation du sol dans la région de Dobrogea :

- les zones rocheuses sont les zones les plus érodées, donc au sommet des collines
- les steppes et les zones de buissons ne sont jamais en plaine
- les villages ne sont pas au sommet des collines
- les roselières sont toujours à proximité d'une zone humide située plutôt en zone de plaine...

Ces observations serviront à définir les fonctions d'appartenances par la suite.

Pour la réalisation de la classification, on tiendra donc compte des 7 canaux de l'image Landsat (présentée précédemment), d'un modèle numérique de terrain SRTM d'une résolution de 90 mètres et également des pentes calculées à partir de ce même MNT. Les relevés terrains seront ajoutés au projet en tant que couche thématique, ce qui permet d'en tenir compte lors de la segmentation, et ainsi éviter le chevauchement d'objets lors de l'import des objets « relevés terrains » (Cf. *Figure 14*).

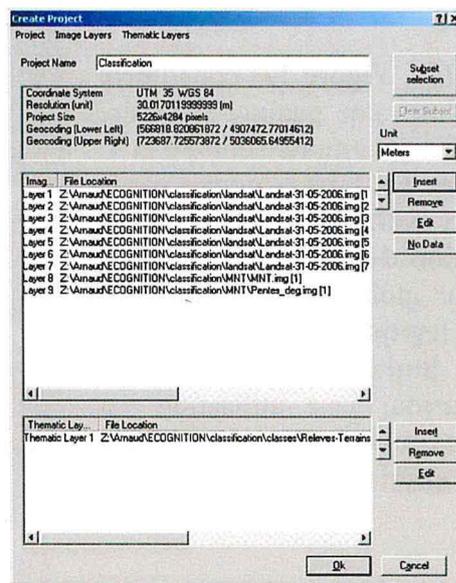


Figure 14 : Création du projet sous eCognition

ii) Segmentation

La première étape consiste à réaliser une segmentation de l'image, afin de la diviser en une multitude d'objets ayant une certaine homogénéité radiométrique. Les paramètres de segmentation sont importants parce qu'ils ont une incidence directe sur la forme et la taille des objets créés et ceux-ci seront la base de tous les traitements.

eCognition permet de travailler à des niveaux de définition différents. Dans notre cas, nous avons choisi de travailler avec un niveau de définition très fin au départ, et de rassembler les objets par la suite. Nous effectuons premièrement une segmentation multi résolution en mode normal avec un paramètre d'échelle de 8 (manque de ressource pour une segmentation plus fine), un facteur de forme faible (0.1). Etant donné que l'on s'intéresse ici aux différences radiométriques des objets, leur forme n'a pas beaucoup d'intérêt, surtout à ce niveau d'échelle. La couche thématique représentant les relevés terrain est utilisée ici dans la segmentation afin qu'il n'y ait pas de chevauchement entre les objets issus des relevés terrain et ceux issus de l'image.

iii) Relevés terrains

Parmi tous les relevés terrains effectués lors de la mission en Roumanie, environ 75% seront utilisés pour réaliser la classification dirigée. Les relevés restants seront utilisés pour l'étape de vérification de la classification. L'ensemble des relevés se présente comme un fichier vecteur de polygones où chaque polygone représente un relevé de végétation (Cf. *Figure 15*). Ce fichier vecteur est ensuite converti en raster pour être chargé dans eCognition en tant que masque TTA (Training and Test Area).

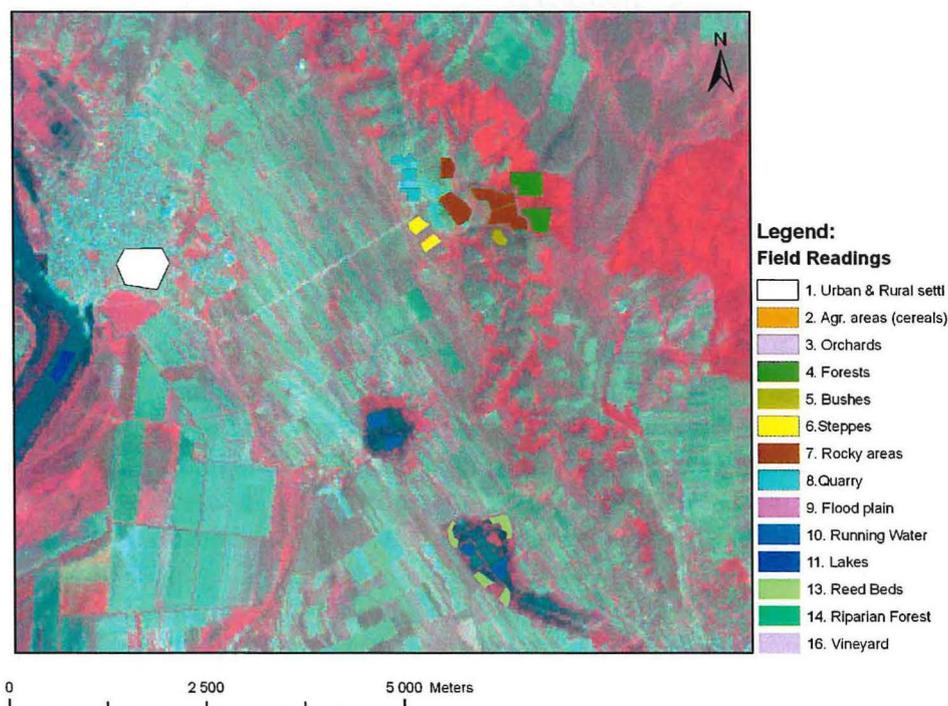


Figure 15 : Extrait des relevés terrains utilisés pour réaliser la classification. En arrière plan : composition colorée à partir de l'image Landsat du 31/05/2006.

iv) Classification

Une fois que ces relevés sont intégrés sous eCognition en tant que TTA masque, et que chaque classe est associée aux différents relevés correspondants, on réalise une classification par plus proche voisin (Nearest Neighbour). Cette classification permet de chercher dans l'image les objets ayant les mêmes propriétés spectrales que les objets donnés en entrée (relevés terrains). Cette méthode est ensuite couplée à l'approche par fonctions d'appartenances qui permettent d'affiner les critères de classification, comme les paramètres d'homogénéité, de forme, de texture ou de localisation par rapport aux autres classes.

Prenons par exemple deux classes similaires radiométriquement comme les ripisylves (forêts le long du Danube) et les forêts du plateau. La grande différence entre ces deux classes est principalement sa localisation et non sa différence spectrale. eCognition permet de tenir compte de la distance spatiale d'une classe par rapport à une autre. Ainsi, seuls les bois à proximité du Danube seront classés comme ripisylves, et les autres pourront être classés comme forêt.

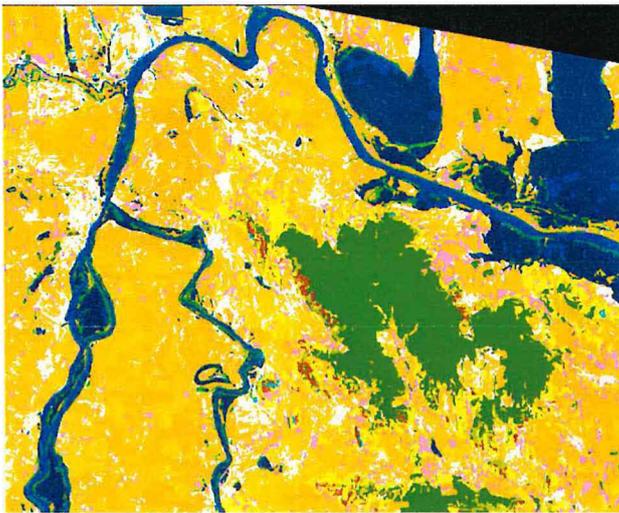


Figure 16 : Extrait de la classification au niveau 1

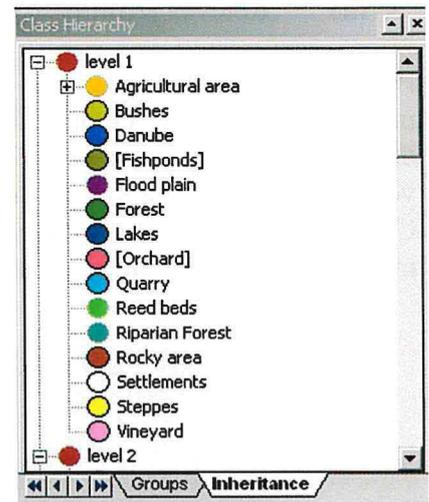


Figure 17 : Hiérarchie des classes pour le niveau 1

Malgré ces paramètres il peut rester de nombreuses confusions entre les classes, et il peut être utile de refaire une nouvelle segmentation, avec un paramètre d'échelle plus élevé en tenant compte de la classification réalisée. Ainsi, les objets auront une taille supérieure mais ils ne pourront pas appartenir à deux classes différentes du premier niveau. Un des avantages de réaliser un nouveau niveau est de pouvoir modifier la hiérarchie et les fonctions d'appartenances sans modifier le travail de classification déjà réalisé, mais il offre également la possibilité de travailler sur des zones précises de l'image et ainsi gagner un temps de calcul très important (ce qui n'est pas négligeable sous eCognition !). Par exemple, si on s'intéresse à l'eau, il suffit de préciser que l'on travaille sur des objets qui étaient classés comme « Danube » ou « Lakes » au niveau 1 et désactiver les autres classes. On peut alors avec les fonctions d'appartenances différencier différentes zones d'eau comme le Danube, les lacs ou la végétation (humide) (Cf. Figure 18).

On constate également qu'au niveau 1 on avait une forte confusion entre les villages et les zones agricoles (sols nus). Le niveau 2 permet entre autre de séparer parmi les objets qui étaient classés en villages ceux qui le sont réellement et ceux qui s'avèrent être du sol nus. Bien que ces deux classes aient des réponses spectrales semblables (forte réponse sur l'ensemble des canaux), les paramètres de texture semblent différencier relativement bien ces

deux classes. On remarque en effet que contrairement aux villages, le sol nu est en général relativement homogène.

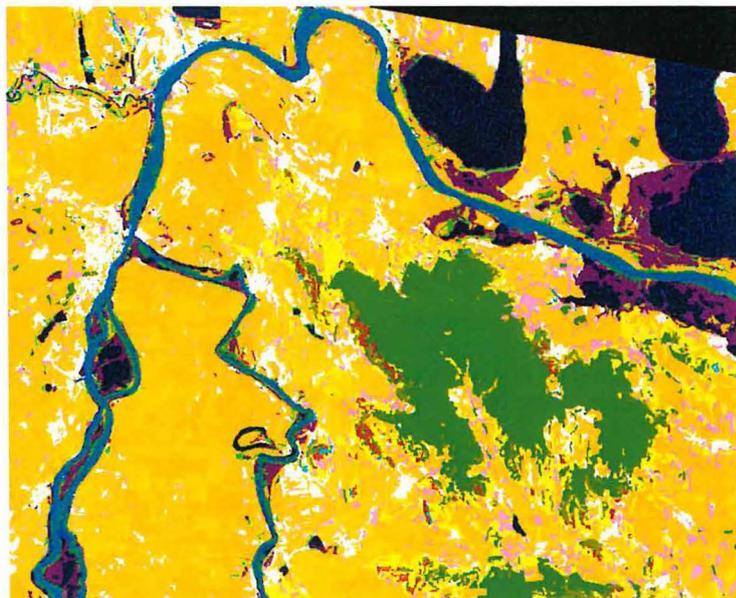


Figure 18 : Extrait de la classification au niveau 2

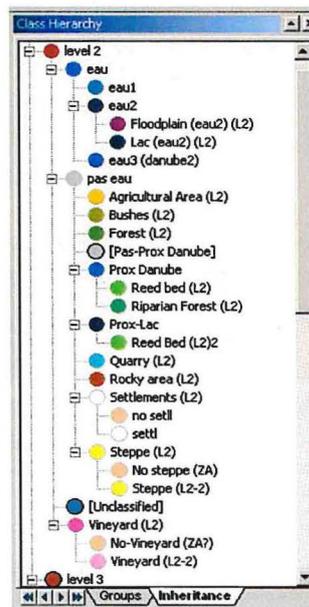


Figure 19 : Hiérarchie des classes au niveau 2

De la même façon que précédemment, un troisième niveau permet d'affiner le deuxième en améliorant la hiérarchie des classes (Cf. Figure 20). Il permet également de travailler sur l'ensemble des objets non classés afin de leur attribuer une classe en fonction de leur réponse radiométrique, de leur altitude et de leur localisation par rapport aux autres classes. En effet si un objet, qui était non classé au niveau 2, est entouré aux trois quarts par une zone agricole, celui-ci a de forte de chance d'appartenir également aux zones agricoles.

Pour finaliser la classification, il est préférable d'utiliser la fonction de fusion des objets sous eCognition qui va permettre de fusionner en un seul objet tous les pixels voisins appartenant à une même classe. Cette simplification de l'image permet d'exporter plus facilement la classification, et cela permet également de la manipuler plus facilement sous un SIG car elle sera beaucoup moins volumineuse.

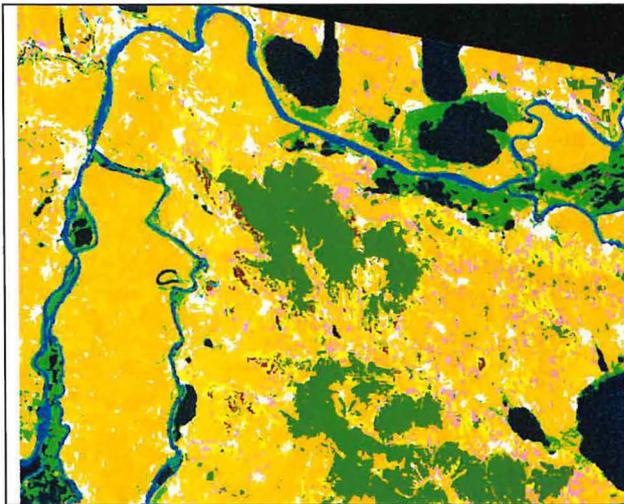


Figure 20 : Extrait de la classification du plateau de Dobrogea au niveau 3

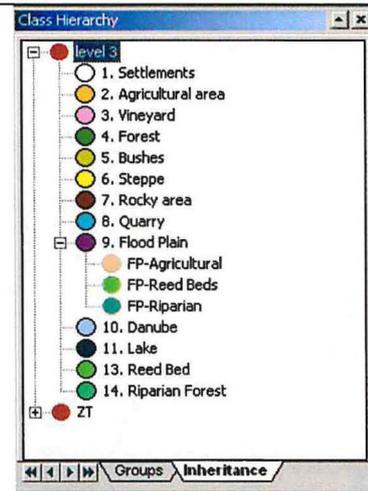


Figure 21 : Hiérarchie des classes au 3è niveau

c) Qualité de la classification

Une fois que la classification a été réalisée, il est important d'évaluer sa qualité afin de voir si elle s'avère fiable ou non. Pour cela une matrice de confusion a été établie. Une partie des relevés terrains ont été sélectionnés pour servir de référence. Chaque classe de la classification possède plusieurs objets test. Le but de la matrice de confusion est de voir la proportion des éléments correctement classés (Cf. Tableau 2). Cette matrice est exprimée en pourcentage correspondant à la proportion d'une classe par rapport au nombre d'objets pour une référence.

Les objets correspondent à la segmentation faite sous eCognition. Etant donné que celle-ci est relativement fine, le résultat est proche de celui d'une matrice exprimée en nombre de pixels.

Tableau 2 : Matrice de confusion de la classification en zone sèche

Classification / Références	Réf. 1	Réf. 2	Réf. 3	Réf. 4	Réf. 5	Réf. 6	Réf. 7	Réf. 8	Réf. 10	Réf. 11	Réf. 13	Réf. 14
1. Zones urbaines	95%	11%	0%	0%	0%	0%	0%	12%	0%	0%	5%	0%
2. Zones agric. (céréales)	4%	75%	7%	3%	0%	19%	1%	0%	0%	1%	21%	33%
3. Vignes	0%	10%	86%	0%	0%	0%	0%	0%	0%	0%	0%	0%
4. Forêts	0%	0%	0%	91%	3%	6%	0%	0%	0%	0%	6%	0%
5. Buissons	0%	2%	0%	0%	92%	2%	8%	0%	0%	0%	0%	0%
6. Steppes	0%	2%	7%	6%	3%	73%	16%	9%	0%	0%	0%	0%
7. Zones rocheuses	0%	0%	0%	0%	3%	0%	75%	4%	0%	0%	0%	0%
8. Carrières	2%	0%	0%	0%	0%	0%	0%	49%	0%	0%	0%	0%
10. Eau courante	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%
11. Lacs	0%	0%	0%	0%	0%	0%	0%	6%	0%	99%	4%	0%
13. Roseaux	0%	0%	0%	0%	0%	0%	0%	21%	0%	0%	65%	0%
14. Ripisylves	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	67%
Nombre d'objets	2119	1731	549	395	231	773	294	191	2865	3464	325	126

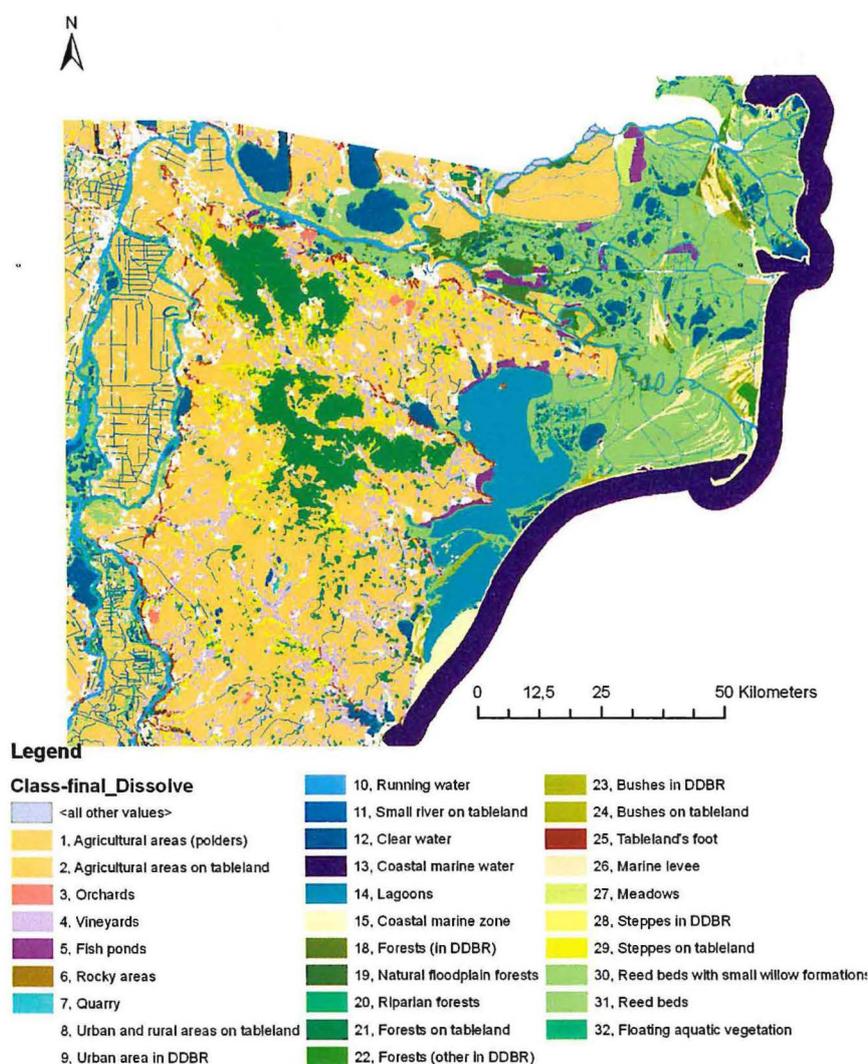
Sur l'ensemble des objets test d'une classe de référence, la matrice montre le nombre d'objets étant correctement classés. Prenons comme exemple la classe « zones urbaines », 95% des objets ont bien classés.

Pour évaluer la classification, on peut calculer la précision totale. Elle correspond à la somme des objets bien classés divisée par le nombre total d'objets. La précision totale de cette classification est de 90% ce qui semble tout à fait satisfaisant.

Le coefficient Kappa permet de faire une évaluation globale de la classification en tenant compte des erreurs d'excédent et de déficit. Son calcul est expliqué en Annexe 5. La classification donne un coefficient kappa de 0,88 ce qui signifie que la classification évite 88% des erreurs obtenues lors d'une affectation au hasard. Ces résultats semblent tout à fait satisfaisants, la tendance est à considérer qu'une bonne classification doit avoir une précision globale (coefficient Kappa) de 85% au minimum [Ref 18].

4) Carte des habitats oiseaux sur l'ensemble de la zone

La carte des habitats oiseaux de l'ensemble de la zone est obtenue par assemblage de la carte des habitats en zone humide et de la classification en zone sèche à laquelle à été superposée la couche rivière et la couche talus. Cette carte permet de visualiser la répartition des différents habitats d'oiseaux dans la région de Dobrogea (Cf. *gure 22*).



Bien que de même nature, certaines classes de la zone sèche (tableland) et de la zone humide (DDBR pour Danube Delta Biosphere Reserve) ont été dissociées. En effet, ces classes sont incluses dans des écosystèmes différents, elles présentent donc des différences en terme de populations d'oiseaux et/ou en terme d'abondance.

gure 22 : Carte des habitats oiseaux sur l'ensemble de la zone d'étude

III) Premières analyses des données ornithologiques et épidémiologiques

La carte des habitats oiseaux couplée à une base de données oiseaux sur la région de Dobrogea par l'intermédiaire d'un SIG constitue un outil précieux pour les ornithologues mais également pour des études écologiques ou épidémiologiques des maladies liées aux oiseaux sauvages comme le virus de West Nile. Nous proposons dans cette partie de présenter les données disponibles pour l'étude, puis les premières analyses des données ornithologiques et finalement de rechercher les facteurs favorables au développement du virus.

1) Les données disponibles

a) La base de données oiseaux

La base de données oiseaux a été remplie par l'équipe ornithologique du delta du Danube, et principalement par Mihai Marinov et Botond Kiss. Cette base de données a été construite sur le modèle de celle effectuée en Camargue dans le cadre du projet EDEN. Elle caractérise l'hétérogénéité, la diversité et l'abondance des oiseaux sauvages sur la zone d'étude [Ref 3]. Au total plus de 300 espèces d'oiseaux ont été recensées dans la zone de Dobrogea, mais seules 199 espèces potentiellement impliquées dans la transmission du Virus de West Nile ont été sélectionnées.

Tableau 3 : Extrait de la base de données oiseaux dans la région de Dobrogea

Species	Season	Activity	Habitat	Habitat index	Occurrence	Occurrence index
Avocet (<i>Recurvirostra avosetta</i>)	spring	feeding	Lagoons	14	common	5
Avocet (<i>Recurvirostra avosetta</i>)	spring	breeding	Lagoons	14	common	5
Avocet (<i>Recurvirostra avosetta</i>)	spring	roosting	Lagoons	14	common	5
Avocet (<i>Recurvirostra avosetta</i>)	spring	feeding	Coastal marine zones	15	abundant	50
Avocet (<i>Recurvirostra avosetta</i>)	spring	breeding	Coastal marine zones	15	abundant	50
Avocet (<i>Recurvirostra avosetta</i>)	spring	roosting	Coastal marine zones	15	abundant	50
Avocet (<i>Recurvirostra avosetta</i>)	spring	feeding	Marshes	17	common	5
Avocet (<i>Recurvirostra avosetta</i>)	spring	breeding	Marshes	17	common	5
Avocet (<i>Recurvirostra avosetta</i>)	spring	roosting	Marshes	17	common	5
Avocet (<i>Recurvirostra avosetta</i>)	summer	feeding	Lagoons	14	common	5
Avocet (<i>Recurvirostra avosetta</i>)	summer	breeding	Lagoons	14	common	5

L'abondance relative de chaque espèce dans chaque unité écologique est évaluée à partir de bibliographie et des connaissances des ornithologistes. L'abondance a été définie comme la probabilité d'occurrence d'une espèce d'oiseau dans une unité écologique pour une saison définie et une activité particulière. Celle-ci a été définie selon 5 classes (Cf. Tableau 4).

Tableau 4 : Definition of classes of probability of occurrence of wild bird species

Class	Probability of occurrence	Definition	Example
0	absent	Never or accidentally observed at a site	Ruff (<i>Philomachus pugnax</i>) in coastalpine forest (!)
1	uncommon	Observed every year, but with only few observations per year	Osprey (<i>Pandion haliaetus</i>) in freshwater marshes
2	frequent	Observed frequently at a site, but not at every visit	Greenshank (<i>Tringa nebularia</i>) in freshwater marshes
3	common	1 to 10 individual observed at a site during a visit	Little egret (<i>Egretta garzetta</i>) in salt marsh lagoon
4	abundant	10 to 100 individual observed at a site during a visit	Stilt (<i>Himantopus himantopus</i>) in salt marsh lagoon
5	very abundant	> 100 individual observed at a site during a visit	Black-headed gull (<i>Larus ridibundus</i>) or flamingo (<i>Phoenicopterus ruber</i>) in salt marsh lagoon

L'indice d'occurrence de la base de données oiseaux (Cf. Tableau 3) a été estimé avec la définition des classes d'occurrence (Cf. Tableau 4). Il représente le nombre d'individus susceptibles d'être observés durant la durée d'une observation dans une unité écologique spécifique [Ref 3].

La construction d'une telle base de données est difficile et demande beaucoup d'observations sur le terrain, notamment pour évaluer l'abondance (vérification par comptage). La base de données en Camargue est toujours en cours de validation après 2 ans de travail. La base de données oiseaux pour la région de Dobrogea n'est vieille que de quelques mois, certaines modifications sont donc susceptibles d'être apportées.

b) Les données sérologiques

Les données sérologiques correspondent aux résultats des prélèvements sérologiques effectués sur les chevaux dans différentes fermes réparties dans toute la région de Dobrogea (Cf. *Figure 23*).

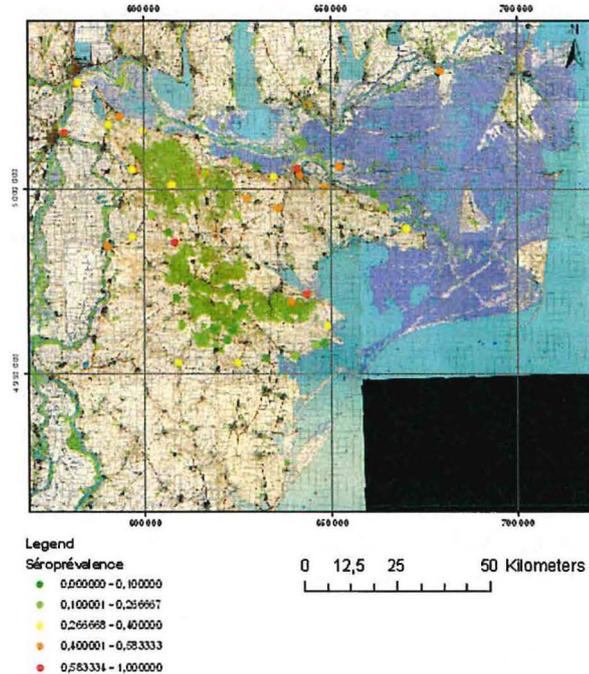


Figure 23 : Localisation des élevages de chevaux ayant fait l'objet de l'étude (extrait d'une carte militaire sur la zone de Dobrogea en fond de carte)

Pour chacune de ces fermes on dispose du nombre de tests effectués, ainsi que le nombre de cas positifs/négatifs trouvés, ce qui nous permet d'en déduire la prévalence (pourcentage du nombre de positifs sur le nombre de tests) (Cf. *Figure 24*).

FID	Shape'	ID	VILLE	IBTEST	IBPOSITIFS	IBNEGATIFS	SERO
0	Multipoint	1	Mircea Voda	6	2	4	0,333333
1	Multipoint	2	Cerna	1	1	0	1,000000
2	Multipoint	3	Casimcea	20	3	17	0,150000
3	Multipoint	4	General Praporgescu	5	1	4	0,200000
4	Multipoint	5	Mircea Voda	2	0	2	0,000000
5	Multipoint	6	Malcoci	10	5	5	0,500000
6	Multipoint	8	Grindu	14	3	11	0,214286
7	Multipoint	9	Parches	15	4	11	0,266667
8	Multipoint	10	Mineri	5	2	3	0,400000
9	Multipoint	11	Jijila	5	1	4	0,200000
10	Multipoint	13	Partizani	26	19	7	0,730769

Figure 24 : Table attributaire des données de séroprévalence

2) Analyse des données ornithologiques

a) Méthodologie

En intégrant la table de données oiseaux dans un logiciel de base de données (Access), on peut facilement créer des requêtes permettant d'interroger la base de données suivant les données voulues. On peut par exemple connaître le nom des espèces ayant une certaine activité dans un habitat choisi, ou connaître le nombre d'espèces présentes pour un type d'habitat et une saison définie, ou alors calculer l'occurrence moyenne des espèces d'oiseaux pour un habitat et une saison données... Ces requêtes peuvent être associées à un SIG par l'intermédiaire d'une liaison OLE afin de les coupler à la carte des habitats oiseaux (par jointure) et ainsi connaître leur distribution spatiale.

Nous proposons dans cette partie d'utiliser cet outil pour tester plusieurs relations écologiques connues. Premièrement, la répartition de la communauté oiseaux peut être observée en effectuant la moyenne des abondances de toutes les espèces d'oiseaux pour chaque habitat. Des cartes représentant le nombre d'espèces d'oiseaux par habitat servent également à visualiser spatialement la diversité des espèces. D'après la bibliographie, l'abondance des espèces est intimement liée au nombre d'espèces présentes dans un écosystème [Ref 14, Ref 15, Ref 16, Ref 17]. Nous proposons d'étudier ensuite l'échelle du paysage par rapport à la population d'oiseaux. En effet, la présence d'un habitat a une influence sur les populations des habitats environnants, et ce jusqu'à une certaine distance. Il est donc important de connaître la distance à partir de laquelle on peut considérer qu'un habitat n'a plus d'influence sur la population oiseaux. Finalement, on étudiera l'impact de l'activité humaine sur la communauté oiseaux.

b) Cartes des abondances moyennes des oiseaux par habitat

Le nombre d'individus d'une espèce pour un habitat, une saison et une activité a été estimé en fonction de la définition de l'indice d'occurrence de la base de donnée (Cf. **Tableau 4**). Les cartes d'abondances correspondent à la moyenne sur l'ensemble des espèces du nombre d'individus susceptibles d'être observés pendant la durée d'une observation pour chaque habitat. Ces cartes donnent ainsi un aperçu de la répartition des oiseaux (toutes espèces confondues) dans la zone d'étude.

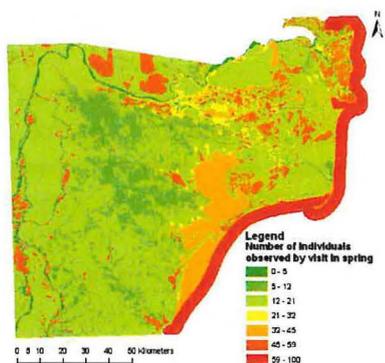


Figure 25 : Carte des abondances moyennes au printemps

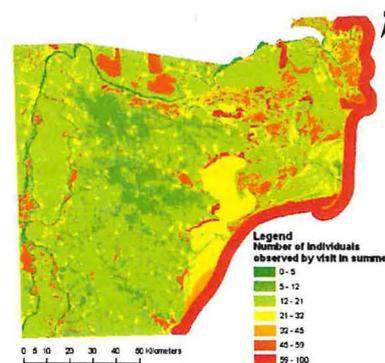


Figure 26 : Carte des abondances moyennes en été

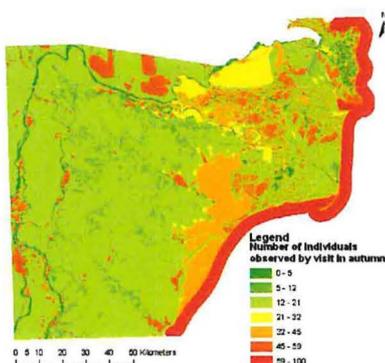


Figure 27 : Carte des abondances moyennes à l'automne

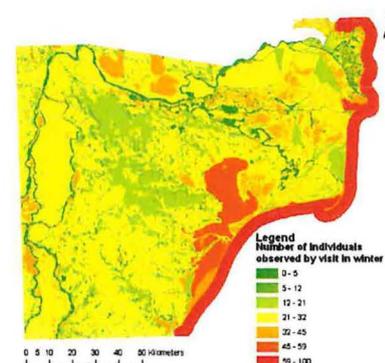


Figure 28 : Carte des abondances moyennes en hiver

c) Cartes du nombre d'espèces d'oiseaux par habitat

D'après la bibliographie, le nombre d'espèces présentes dans un écosystème est une caractéristique de l'écosystème. Il témoigne de la diversité des espèces et permet ainsi de définir si le milieu est plus ou moins naturel (activité humaine) [Ref 14, Ref 15, Ref 16, Ref 17]. Il nous a semblé intéressant de compter le nombre d'espèces d'oiseaux présentes pour chaque habitat et de le cartographier. Les résultats sont assez surprenants car ils montrent que d'après la base de données remplie par l'équipe ornithologique du delta du Danube, il y a une plus grande diversité des espèces dans la zone sèche (plateau Dobrogea) que dans la zone humide (le delta) (Cf. **Figure 29** à **Figure 32**). Ces résultats doivent être discutés avec l'équipe ornithologique et doivent être pris avec précaution pour la suite.

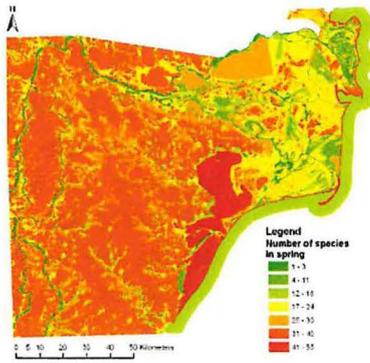


Figure 29 : Carte du nombre d'espèces au printemps

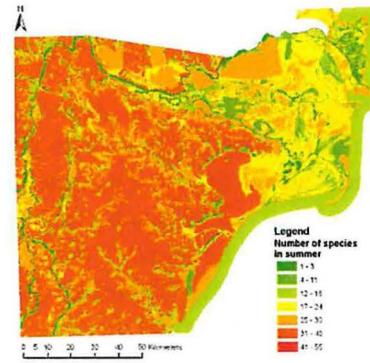


Figure 30 : Carte du nombre d'espèces en été

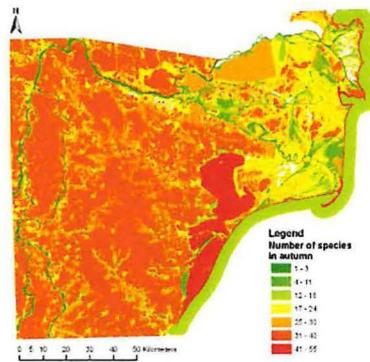


Figure 31 : Carte du nombre d'espèces à l'automne

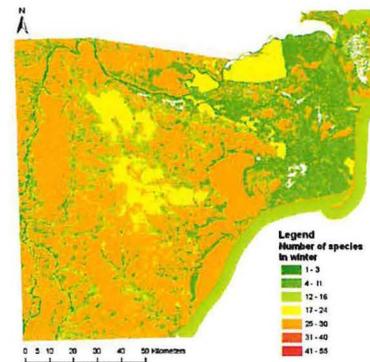


Figure 32 : Carte du nombre d'espèces en hiver

d) Echelle de paysage et communauté oiseaux

La répartition des oiseaux a été décrite à l'aide d'une carte des habitats où chaque habitat a une caractéristique propre en matière de population et d'abondance. Cependant il est évident que les oiseaux se déplacent et que la présence d'un habitat dans une région donnée a une influence sur les populations au-delà des limites de celui-ci.

Pour connaître l'échelle de paysage à utiliser dans l'analyse du voisinage des élevages sains et infectés (Cf. III) 3)) nous avons calculé plusieurs variogrammes : sur l'ensemble de la zone, sur la zone sèche, et sur la zone humide pour les 4 saisons. La variable utilisée pour le calcul est le nombre d'espèces d'oiseaux présentes. En effet, chaque point est lié à un habitat qui est associé à un certain nombre d'espèces. Le variogramme permet de visualiser la variation du nombre d'espèces dans l'espace.

Un variogramme permet de mesurer la variabilité (ou la dissemblance) à différentes échelles d'une variable régionalisée ($Z(x)$). Il est défini par l'expression suivante (Cf. Équation 1) [Ref 20], où E correspond à l'espérance mathématique et h à la distance séparant un couple de points.

$$\gamma(h) = \frac{1}{2} \cdot E[(Z(s) - Z(s+h))^2]$$

Équation 1 : Formule du (semi) variogramme

Le variogramme est une fonction croissante, bornée, nulle en zéro. A partir d'une certaine distance, le variogramme atteint le plateau, on dit que les données sont décorréélées à partir de cette distance. Si le plateau est atteint pour une distance courte, cela implique que les données sont peu corrélées. En général, on peut considérer que les données sont corrélées tant que le variogramme est en dessous d'un certain seuil. Si ce seuil est dépassé, ou que le résultat est proche de ce seuil cela implique que les données spatiales sont semblables à des données aléatoires, il n'y a donc plus de corrélation.

Les résultats de variogrammes ont été difficiles à interpréter car la carte semble toujours corrélée même pour des distances importantes de l'ordre de 50 km. On remarque tout de même que les variogrammes de la zone sèche semblent atteindre la valeur de seuil à partir de 2 à 3 km selon la saison (Cf. *Figure 33*). Ce résultat signifie qu'au delà de 2 à 3 km, l'occupation du sol n'a plus d'influence sur la communauté oiseaux. L'échelle de paysage retenue pour la suite des analyses sera donc de 2 km.

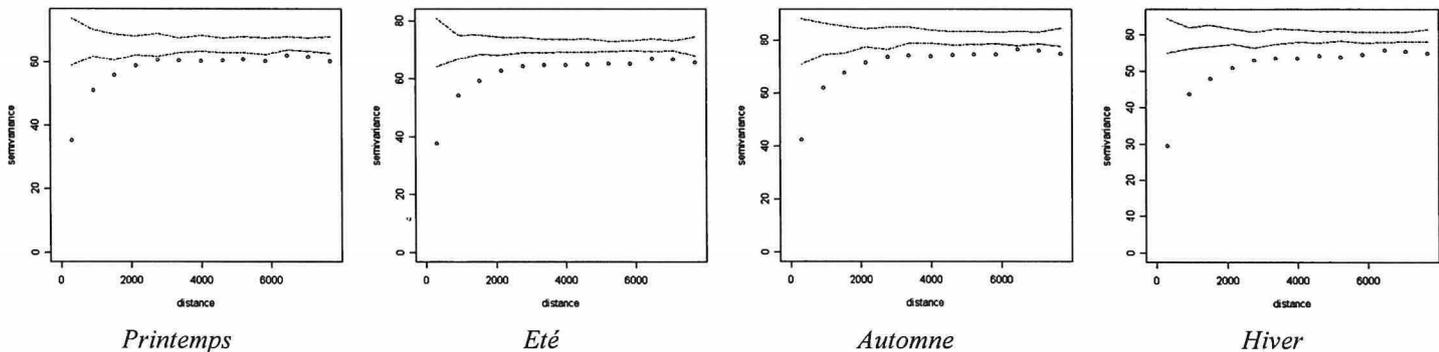


Figure 33 : Variogrammes sur la zone sèche pour les 4 saisons

e) Impact de l'urbanisation sur la communauté oiseaux

Pour évaluer l'impact de l'activité humaine, nous avons tracé un graphe du nombre d'espèces d'oiseaux en fonction d'un taux d'activité humaine. Deux méthodes différentes ont été utilisées. La première consiste à associer à un certain nombre de points dont on connaît le nombre d'espèces présentes à la distance aux zones d'activités humaines. La deuxième méthode consiste à attribuer un indice d'activité humaine pour chaque habitat qui sera associé à leur nombre d'espèces présentes.

i) Distance aux zones de forte activité humaine

Nous avons choisi de prendre comme zones à forte activité humaine l'ensemble des zones urbaines ainsi que la classe eau courante qui est constitué principalement d'axes maritimes importants. Une carte de la distance euclidienne à ces zones à été créée à l'aide de l'extension « spatial analyst » sous ArcGis (ESRI). Un nuage de points aléatoires sur la zone d'étude a été associé au nombre d'espèces par saison (par jointure spatiale) avec la carte du nombre d'espèces présentée précédemment. Ce nuage de points à ensuite été projeté sur la carte des

distances euclidiennes aux zones à forte activité humaine, afin d'obtenir pour chacun des points l'information sur le nombre d'espèces et sur l'indice d'urbanisation.

Cette méthode a permis de tracer les graphes du nombre d'espèces présentes en fonction de la distance aux zones d'activités humaines pour les 4 saisons (Cf. Figure 34). On suppose que les points dont la distance est supérieure à 4 km sont principalement des points à l'intérieur de la zone humide car ils nécessitent l'absence de toute zone urbaine dans un rayon de 4 km. En revanche, les points dont la distance est inférieure à 4 km ont plus de chance d'appartenir à la zone sèche. Les courbes sont assez bruitées mais on a vu précédemment (Cf. III) 2) c)) que la diversité des espèces entre la zone humide (moins anthropisée mais avec un nombre inférieur d'espèces) et la zone sèche (davantage anthropisée avec un plus grand nombre d'espèces) devait être discutée et vérifiée par les ornithologues.

Si l'on considère uniquement les points de la zone sèche et que l'on tient compte de l'échelle d'analyse du paysage de 2 km, les courbes de la Figure 34 semblent montrer une nette augmentation de la diversité des espèces avec la distance aux à forte activité humaine, ce qui correspond tout à fait au résultat attendu.

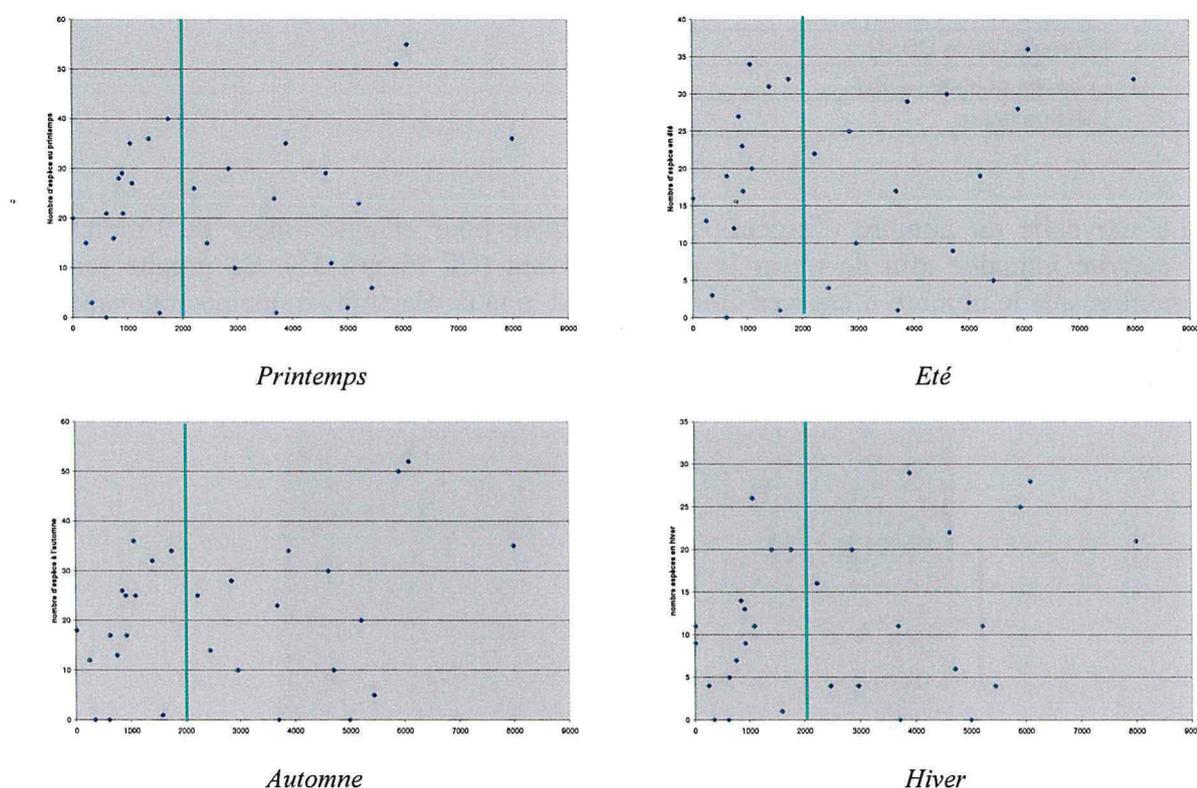


Figure 34 : Impact urbanisation sur la communauté oiseaux. Nombre d'espèces en fonction de la distance aux zones à forte activité humaine exprimée en mètres

ii) Indice d'urbanisation pour chaque habitat

Un indice de l'activité humaine sur une échelle de 1 à 5 a été attribué à chaque classe. Un indice de 5 a été attribué aux zones urbaines et aux grands axes fluviaux, tandis que les zones isolées du delta ont un indice de 1. Les zones de cultures quant à elles ont un indice de 4 car on considère qu'elles présentent une activité humaine régulière (Cf. Tableau 5).

Tableau 5 : Attribution du d'un indice d'activité humaine par habitat

Human activity index	Habitat	
1	Rocky areas on Tableland	
	Coastal marine waters	
	Lagoons	
	Forests (mixed oak wood on marine levee)	
	Natural flood plain forest	
	Forests	
	Bushes	
	Meadows	
	Reed beds	
	Flating Aquatic Vegetation (Clairs de roselières)	
	2	Clear water
		Coastal marine zone
Forest outside DDBR		
Bushes outside DDBR		
Marine levee		
3	Agricultural areas	
	Agricultural areas outside DDBR	
	Orchards	
	Vineyards	
	Small rivers on Tableland	
	Tableland's foot (undercliffs)	
4	Steppes on Tableland	
	Fish ponds	
	Quarry	
5	Riparian forest	
	Urban and rural settlements outside DDBR	
	Urban and rural settlements	
	Running waters	

La moyenne du nombre d'espèces a été calculée pour chaque valeur de l'indice de l'activité humaine afin de tracer le graphe suivant (Cf. Figure 35). Ce graphe semble montrer que le nombre d'espèces diminue quand le taux d'activité humaine augmente, ce qui semble tout à fait cohérent.

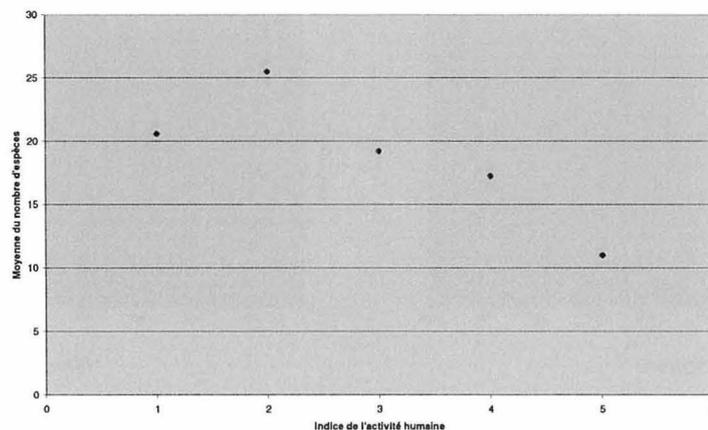


Figure 35 : Nombre d'espèces présentes en fonction de l'indice de l'activité humaine

3) Recherche de facteurs de risque West Nile

La recherche de facteurs de risque pour le virus de West Nile consiste à associer les données de séroprévalence avec des données environnementales propres à chaque ferme comme l'occupation du sol, l'abondance moyenne et le nombre d'espèces d'oiseaux à proximité de chaque ferme et à déterminer les paramètres corrélés aux données de prévalence (analyse univariée). En réalisant des régressions linéaires multiples (analyse multi-variables) on pourra construire des modèles permettant de prédire la prévalence avec un degré de fiabilité et une marge d'erreur connue. L'étape suivante serait d'estimer la prévalence sur l'ensemble de la zone en fonction des paramètres environnementaux et de cartographier le risque lié à la prévalence. Cette dernière étape ne sera pas réalisée car les données disponibles ne semblent pas suffisantes pour décrire la transmission du virus, d'autres paramètres environnementaux devront être testés et une analyse statistique plus poussée doit être conduite pour pouvoir construire un modèle permettant de prédire la prévalence avec une bonne précision.

L'échelle paysagère choisie pour la réalisation de l'étude est de 2 km, correspondant à l'échelle définie à l'aide des variogrammes (Cf. III) 2) d)). Pour chaque ferme, le pourcentage des habitats présents dans un rayon de 2 km est calculé. Ce calcul permet d'établir un tableau dans lequel figure, pour chaque ferme, la séroprévalence, l'occupation du sol, l'abondance moyenne des oiseaux et le nombre d'espèces.

Pour réaliser ce travail, nous avons utilisé le logiciel R, qui est un langage de programmation et un environnement mathématique utilisés pour l'analyse statistique. R est un logiciel libre qui propose des centaines de bibliothèques directement téléchargeables sur internet.

a) Analyses des variables

Au total nous disposons de 40 variables pour décrire les données de séroprévalence, correspondant aux 32 habitats oiseaux, 4 valeurs d'abondance moyennes et 4 valeurs du nombre d'espèces présentes autour des fermes (pour chacune des saisons). L'analyse univariée permet de calculer une matrice de corrélation pour connaître les ressemblances des variables entre elles et ainsi sélectionner les variables les plus pertinentes.

Pour connaître les coefficients de corrélation entre les données de séroprévalence et chacune des variables, nous utilisons la fonction « cor.test » du logiciel R qui permet de tester la corrélation entre deux variables et qui fournit la probabilité critique du test (p-value). Cette dernière témoigne de la pertinence du couple de données dans le calcul. Elle doit être inférieure à 10% pour que la variable testée soit considérée comme significative.

Au final, très peu de variables s'avèrent être significatives, seule les unités zones urbaines en zone humide, zones urbaines en zone sèche et occurrence moyenne en été ont une p-value inférieure à 10% (Cf. Tableau 6).

Tableau 6 : Coefficients de corrélation et p-values par rapport aux données de séroprévalence

Variables	coefficient de corrélation	p-value
Zones urbaines en zone humide	0,338	0,017
Zones urbaines en zones sèches	0,254	0,076
Occurrence moyenne en été	0,236	0,098

Pour essayer d'améliorer les résultats, des regroupements de classes sont effectués pour ne garder que 8 grandes unités :

- Une unité agricole regroupant les classes de cultures céréalières (delta et zone sèche), les vergers et les vignes.
- Une unité eau regroupant les lacs, lagunes, l'eau courante et les ripisylves.
- Une unité Zones rocheuses regroupant les zones rocheuses et les carrières
- Une unité zones urbaines (delta et plateau)
- Une unité forêts
- Une unité steppes regroupant les prairies, steppes et buisson
- Une unité roseaux
- Une unité occurrence correspondant à la moyenne des abondances des oiseaux en été, à proximité de chaque ferme.

Les coefficients de corrélations (R^2) et les p-values sont calculés pour chacune de ces unités (Cf. Tableau 7). Le regroupement s'avère être positif car certaines classes qui ne semblaient pas être significatives au départ ont une certaine ressemblance avec les données de séroprévalence une fois regroupées. C'est le cas pour les unités eau et forêts qui seront sélectionnées avec les unités zones urbaines et eau pour la suite de l'étude. Les unités agricoles, zones rocheuses, buissons, steppes et roseaux ne sont pas significatives (p-value supérieur à 10%) elles ne seront donc pas exploitables pour la suite du traitement.

Le coefficient de corrélation montre que le risque d'infection augmente en priorité avec la présence de zones urbaines. Un deuxième facteur aggravant est la présence d'eau, suivi d'une abondance moyenne élevée des oiseaux. En effet les valeurs du coefficient sont positives et comprises entre 25 et 38%. La forêt est au contraire une classe défavorable au virus, son coefficient de corrélation est négatif. La forêt peut alors être considéré comme un facteur protecteur.

Tableau 7 : Coefficients de corrélation et p-values par rapport aux données de séroprévalence

Variables	coefficient de corrélation	p-value
Agricole	-0,043	0,767
Eau	0,272	0,056
Zones rocheuses	0,030	0,834
Urbain	0,385	0,006
Forêts	-0,247	0,084
Buissons	-0,059	0,685
Steppes	-0,075	0,603
Roseaux	0,099	0,493
Occurrence	0,236	0,098

b) Analyse multi-variables

A partir des variables significatives sélectionnées précédemment (eau, zones urbaines, forêt et occurrence), plusieurs modèles vont être créés. Ils correspondent à l'association de plusieurs variables. Chacun de ces modèles est obtenu par un calcul de régression linéaire multiple à l'aide de la fonction « lm » sous le logiciel R. Au total 11 modèles différents seront testés (Cf. la liste des modèles dans le Tableau 8). Le but de ces modèles est de prédire la prévalence en fonction de ses paramètres environnementaux. Chacun de ces modèle présente des imprécisions, il est donc important de les évaluer pour connaître le ou lesquels présentent les meilleurs résultats.

Nous avons utilisé le critère d'information d'Akaike pour discriminer ces modèles. Ce critère s'applique aux modèles estimés par régression linéaire multiple, il permet de les évaluer en tenant compte de leur nombre de paramètres. Etant donné que l'on dispose d'un nombre limité d'observations (n=50 fermes) par rapport au nombre de paramètres (k compris entre 3 et 6), il est préférable d'utiliser le critère d'information d'Akaike corrigé (AICc) (n/k<40) [Ref 19] (Cf. équation du critère d'Akaike corrigé en Annexe 4).

Le critère d'Akaike est calculé pour chacun des modèles, ce qui permet de les classer du plus significatifs au moins fiable, le meilleur étant celui dont le critère d'Akaike corrigé est le plus faible (Cf. Tableau 8).

Tableau 8 : Classement de différents modèles suivant le critère d'information d'Akaike corrigé (AICc)

modèle	K	n	AIC	AICc
Urbain + Eau	4	50	-7,3	-6,41
Urbain	3	50	-6,86	-6,33
Urbain + Occurrence	4	50	-5,26	-4,37
Urbain + Forêt	4	50	-5,16	-4,26
Urbain + Eau + Forêts	5	50	-5,45	-4,08
Urbain+Eau+Occurrence	5	50	-5,43	-4,06
Eau	3	50	-2,67	-2,15
Eau + Forêt	4	50	-2,85	-1,96
Urbain+Eau+Forêts+Occurrence	6	50	-3,64	-1,68
Forêts	3	50	-1,97	-1,44
Occurrence	3	50	-1,7	-1,18

Le meilleur modèle est celui faisant intervenir les zones urbaines et l'eau. Le résultat de la régression linéaire multiple (Cf. Tableau 9) montre que les zones urbaines ont une importance plus grande avec un indice de fiabilité (p-value) de 0,01. La prévalence peut être estimée par l'équation suivante :

$$Pr_{ev} = 0,19 + 0,69 \cdot URB + 0,51 \cdot EAU$$

Tableau 9 : Résultat de la régression linéaire multiple pour le modèle "Zones urbaines + Eau"

Variable	Coefficient de régression	Ecart-type	t-value	p-value
Constante	0,19	0,05	3,58	0,00
Zones urbaines (URB)	0,69	0,27	2,58	0,01
Eau	0,51	0,33	1,54	0,13

Cette première analyse permet donc d'identifier la présence de zones urbaines et la présence d'eau comme facteurs de risque pour la circulation du virus de West Nile. Chez les chevaux. L'urbanisation étant inversement liée aux nombres d'espèces d'oiseaux présentes (Cf. III) 2) e)), ce résultat préliminaire suggère que l'émergence du virus de West Nile chez les équidés n'est pas liée à une plus forte diversité d'oiseaux mais à la présence d'espèces d'oiseaux associées au milieu urbain.

L'étape suivante serait d'estimer la prévalence sur l'ensemble de la zone en fonction des paramètres environnementaux et de cartographier le risque lié à la prévalence. Cette dernière étape ne sera pas réalisée ici car les données disponibles ne sont pas suffisantes pour décrire la transmission du virus, d'autres paramètres environnementaux devront être testés pour essayer de construire un modèle statistique permettant de prédire la prévalence avec une bonne précision. Ce travail sera effectué dans le cadre du projet EDEN en parallèle d'une approche de modélisation mathématique du cycle complet de transmission du virus de West Nile (collaboration CIRAD – IRD – Institut du Danube).

IV) Discussion et conclusion

a) Bilan du travail réalisé

Le travail réalisé pendant ce stage est une étape importante dans le cadre de l'étude de l'analyse paysagère de la circulation du virus de West Nile réalisée dans le cadre du projet EDEN. Il permettra d'étudier de manière plus précise et plus complète les conditions environnementales favorables au virus et s'avère également une étape indispensable pour mener une étude comparative avec les autres sites européens.

L'outil développé au cours de ce stage permet de cartographier chaque espèce d'oiseaux susceptible de jouer un rôle dans la transmission de West Nile en précisant son abondance et son activité selon une saison donnée. Il s'avère également très utile pour l'équipe ornithologique en permettant de localiser chaque espèce souhaitée. Cet outil leur permettra certainement d'améliorer leur base de données ou de corriger d'éventuelles incohérences. Les modifications engendrées seront facilement intégrables au SIG mis en place.

Finalement l'outil développé a permis de caractériser la communauté oiseaux et identifier les premiers facteurs de risques du virus de West Nile.

b) Evolution et perspectives

La base de données oiseaux réalisée par l'équipe épidémiologique du Danube Delta National Institute constitue une première étape qui a permis de réaliser mon travail, et permettra de commencer à rechercher des facteurs de risque pour le virus de West Nile. Cette base de données est encore relativement jeune (quelques mois), si on la compare à la base de données en Camargue qui est encore en cours de validation après 2 ans de travail. L'outil développé durant mon stage sera un excellent moyen pour vérifier la base de données oiseaux et corriger d'éventuelles incohérences. Ces améliorations pourraient apporter des modifications dans les listes des habitats mais les changements seront facilement intégrables dans un SIG.

Après la caractérisation de la répartition des hôtes du virus de West Nile, un travail similaire sur les vecteurs devra être réalisé. En effet, connaître la répartition des moustiques permettrait de localiser les zones de contacts hôtes-vecteurs et ainsi localiser les principales zones à risques. Les moustiques ayant un rayon d'action réduit de l'ordre du kilomètre, et ayant une durée de vie plus courte, il s'avère être un élément très important pour le suivi du virus et la localisation des zones à risque. Pour le moment les données disponibles ne sont pas suffisantes pour être interpolée sur la zone entière. Il faudrait pour cela réaliser des captures sur l'ensemble de la zone pour chaque habitat entomologique.

La recherche des facteurs favorable au virus de West Nile a été effectuée avec l'occupation du sol et la moyenne de l'abondance des espèces en été comme variables explicatives. On a remarqué (Cf. III) 4) b)) que certains de ces paramètres avaient une influence sur l'éventuelle présence du virus, cependant elles ne sont pas suffisantes pour caractériser, à elles seules, la transmission du virus de West Nile aux chevaux. Pour une étude complète sur les facteurs de risques il faudra tenir compte de paramètres environnementaux supplémentaires comme la température, l'hygrométrie et des données entomologiques.

L'occurrence moyenne des espèces d'oiseaux en été semble corrélée avec les données de séroprévalence. Pour améliorer ces résultats, on peut très bien imaginer ne sélectionner les espèces d'oiseaux reconnues comme étant impliquées dans la transmission du virus dans d'autres sites avec une activité favorable (repos). Cette sélection permettra de réaliser des

cartes d'abondances différentes et vraisemblablement davantage corrélées avec les données de séroprévalences.

Le troisième élément qui pourrait améliorer la recherche de facteurs de risque de West Nile serait d'utiliser davantage de données épidémiologiques. En effet, on peut utiliser les données de séroprévalences sur les chevaux à plusieurs dates de l'année, ou sur plusieurs années. On peut également envisager d'utiliser les données issues des captures oiseaux et moustiques sur le terrain (ces données n'étaient pas disponibles au moment du stage).

Finalement, la réalisation d'un modèle complet regroupant les données environnementales les plus pertinentes pourront par la suite permettre de réaliser un modèle afin de connaître les éléments les plus favorables au virus et construire des cartes de risque.

c) Conclusion

L'outil développé pendant mon stage est la première étape indispensable à l'étude des conditions favorables à West Nile. Il va permettre non seulement d'outil de comparaison pour les données des différents sites du projet EDEN, mais il permettra également de faire une étude plus poussée sur la recherche facteurs favorables, sur la réalisation de cartes de risques et sur leur prévision qui est un des objectifs du projet EDEN.

D'un point de vue personnel ce stage m'a permis d'acquérir une certaine expérience dans le domaine de la télédétection et de la géomatique pour des applications épidémiologique. Il m'a également permis de me familiariser avec le domaine de la recherche de risque (régression) et donc l'utilisation de quelques outils de géostatistique qui est souvent utilisé quand les données de télédétection servent à caractériser des données environnementales. Cependant, la plus grande richesse de ce stage a certainement été de travailler avec des personnes d'origines scientifiques complètement différentes (ornithologue, entomologiste, vétérinaire, épidémiologiste et spécialiste en télédétection et géomatique) permettant sans cesse d'apprendre des notions nouvelles dans tous ces domaines.

Bibliographie :

- Ref 1 A. Tran, N. Gaidet, G. L'Ambert, T. Balenghien, G. Balança, V. Chevalier, V. Soti, C. Ivanès, E. Etter, F. Schaffner, T. Baldet, S. De la Rocque.
On the use of Remote Sensing for the Ecological Description of Multi-host Disease systems: A case study on West Nile Virus in Southern France.
- Ref 2 A. Leblond, A. Sandoz, G. Lefebvre, H. Zeller, D.J. Bicout
Remote Sensing based identification of environmental risk factors associated with West Nile disease in horses in Camargue, France
ScienceDirect
- Ref 3 N Gaidet, A Tran, G Balança CIRAD, Montpellier, 12 mai 2006
Eco-Epidemiological Study of West Nile virus, rapport EDEN-WNV
- Ref 4 Maria A. DIUK-WASSER, Heuidi E. BROWN, Theodore G. ANDREASDIS,
Durland FISH
Modeling the spatial Distribution of Mosquito Vectors for West Nile Virus in Connecticut, USA
Vector Borne and Zoonotic Diseases - Vol 6, Number 3, 2006
- Ref 5 Vanessa O. Ezenwa, Marwin S. Godsey, Raymond J. King and Stephen C. Gurtill
Avian Diversity and West Nile virus : testing associations between biodiversity and infectious disease risk
The Royal Society – 18 April 2005
- Ref 6 E. Jourdain, Y. Toussaint, A. Leblond, D.J. Bicout, P. Sabatier and M. Gauthier-Clerc
Bird Species Potentially Involved in Introduction, Amplification, and Spread of West Nile Virus in a Mediterranean Wetland, the Camargue (Southern France)
Vector-Borne and Zoonotic Diseases - Volume 7, Number 1, 2007
- Ref 7 Thomas Balenghien : De l'identification des vecteurs du virus de West Nile à la modélisation du risque d'infection dans le sud de la France.
Thèse de doctorat, Ecole doctorale de l'ingénierie pour la santé, la Cognition et l'environnement, soutenue le 20 décembre 2006, 235 pages
- Ref 8 Ezenwa V.O., Godsey M.S., King R.J. & Gurtill S.C. 2006. Avian diversity and West Nile virus: testing associations between biodiversity and infectious disease risk. Proceedings of the Royal Society B-Biological Sciences, 273, 109-117.
- Ref 9 Hay S.I. 2000. An overview of remote sensing and geodesy for epidemiology and public health application.
Advances in parasitology, 47, 1-35.
- Ref 10 Huh O.K. & Malone J.B. 2001. New tools: potential medical applications of data from new and old environmental satellites.
Acta Tropica, 79, 35-4.
- Ref 11 Beck L.R., Lobitz B.M. & Wood B.L. 2000. Remote Sensing and Human Health : new sensors and new opportunities.
Emerging infectious diseases, 6, 217-226.

- 
- Ref 12 Vegetation of the Biosphere Reserve “Danube Delta”
Ministry of Waters and Environment Protection
Danube Delta Biosphere Reserve Authority
National Academy of Sciences of Ukraine
The Ukrainian Danube Delta Biosphere Reserve
M.G.Kholodny Institute of Botany
- Ref 13 Ecosystems of the Romanian Danube Delta Biosphere Reserve
Petre Gastescu, Mircea Oltean, Iulian Nichersu, Adrian Constantinescu
- Ref 14 Graham Bell Neutral Macroecology
Science , Vol 293, 28 september 2001
- Ref 15 Jérôme Chave - Neutral theory and community ecology
(Laboratoire Evolution et Diversité Biologique, CNRS and Université Paul Sabatier)
Ecology Letters 2004
- Ref 16 R.G. Huges - Theories and Models of Species Abundance
The American Naturalist, Vol 128, No 6 Dec 1986
- Ref 17 F.W. Preston - The Comonees, and rarity, of species
Ecology, vol 29, No 3 – Juillet 1948
- Ref 18 S. Durrieu – La validation des classes en Télédétection
Cours du Mastère SILAT 2007
- Ref 19 R. LANCELOT et M. LESNOFF - Sélection de modèles avec AIC et critères
d’information dérivés
Document Cirad
ftp://ftp.cirad.fr/pub/group-r/groupe-r/Fiches/AIC_v3.pdf
- Ref 20 Michel Arnaud et Xavier Emery – Estimation et interpolation spatiale
Hermes Science, sortie en septembre 2000
- Ref 21 John Harte, Ann Kinzig, Jessica Green - Self-Similarity in the Distribution and
Abundance of Species
SCIENCE - VOL 284 - 9 APRIL 1999
- 
- 

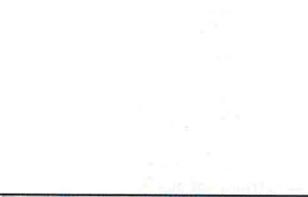
Annexes

Annexe 1 : Liste des habitats oiseaux

CODE	Ecological unit	Field picture	Description
1	Agricultural areas (polders)		<i>Cereals, or unused lands</i>
2	Agricultural Area on Tableland (cereals)		<i>cereals</i>
3	Orchard on Tableland		<i>peaches</i>
4	Vineyard		-
5	Fishponds		<i>Wet or dry</i>
6	Rocky Area on Tableland		<i>(on top of hills), equisetum arvense, mentha arvensis</i>
7	Quarry		-

8	Urban and rural settlement on Tableland		-
9	Urban and rural settlement in DDBR		-
10	Running water		<i>Danube, rivers and channels</i>
11	Small river on Tableland		-
12	Clear Water		-
13	Coastal marine water		-
14	Lagoon		-
15	Coastal marine zone		<i>Crambe maritima, Eryngium maritimum Petasites spurius, Elymus giganteus</i>
16	Fresh water marshes		-

17	Marshes		<i>Schoenoplectus Lacustris, Salicornia patula, Spergularia media, Tamarix ramosissima...</i>
18	Forest (mixed oak wood on marine levee)		<i>Quercus pedunculiflora, Quercus robur, Fraxinus angustifolia, Alnus glutinosa</i>
19	Natural floodplain forest		<i>Salix alba, salix frgilis</i>
20	Riparian forest		<i>Alnus glutinosa, Ashes Fraxinus angustifolia, Salix alba, salix frgilis, Periploca graeca, Viburnum opulus, Populus canadensis</i>
21	Forest on tableland		<i>Quercus pedunculiflora, Q. pubescens, Q. robur, Tilia tomentosa, Fraxinus ornus, Fraxinus ornus, F. excelsior, Pyrus salicifolia</i>
22	Forest (other)		<i>Alnus glutinosa, Fraxinus angustifolia</i>
23	Bushes		<i>Tamarix ramosissima, Elaeagnus Angustifolia, Secale silvestre, and Elymus giganteus</i>
24	Bushes on tableland		<i>Carpinus orientalis,</i>

25	Tableland's foot		<i>Undercliffs 1 to 5m high</i>
26	Marine levee		<i>Festuca beckeri, Carex colchica, Ephedra distachya, Secale silvestre, and Elymus giganteus</i>
27	Meadows		<i>Agrostis stolonifera, Aeluropus littoralis, Puccinellia gigantea, Tripolium vulgare</i>
28	Steppes		<i>Carex colchica, Ephedra distachya, Secale silvestre, Elymus giganteus and Festuca beckeri</i>
29	Steppe on Tableland		<i>Carex colchica, Ephedra distachya, Secale silvestre, Elymus giganteus and Festuca beckeri</i>
30	Reed beds (with small willow formations)		<i>Typha angustifolia, acustris, Apium nodiflorum</i>
31	Reed beds (other)		<i>Typha angustifolia, Scirpus lacustris, Apium nodiflorum</i>
32	Floating aquatic vegetation (Clairs de roselière)		<i>Lemna minor, Salvinia natans, charion fragilis, Nymphoidetum peltatae, Trapetum natantis, Potametum natantis</i>

Annexe 2 : Liste des habitats et association avec la carte de végétation

Bird Data Base Legend		Origine	Vegetation Map Legend	
No	Description		No	Description
1	Agricultural areas (polders)	VM	45	Agricultural polder
2	Agricultural Area on Tableland (cereals)	CL		
3	Orchard on Table land	CL		
4	Vineyard	CL		
5	Fishponds	VM	46	Fish ponds
6	Rocky Area on Tableland	CL		
7	Quarry	CL		
8	Urban and rural settlement on Tableland	CL		
9	Urban and rural settlement in DDBR	VM	47	Urban and rural settlements
10	Running water	VM	44	Algae sp, turbid water
		SP		
		CL		
11	Small river on Tableland	SP		
12	Clear Water	VM	41	Floating aquatic vegetation
		VM	42	Submerged aquatic vegetation
		VM	43	Sparse macrophytes (clear water)
		VM	44	Algae sp, turbid water
		CL		
13	Coastal marine water	SP		buffer along the coast (in the sea) (5km)
14	Lagoon	VM	44	Algae or turbid water (not all)
15	Coastal marine zone	VM	1	Seashore pioneer vegetation
		SP		buffer along the coast (inland) (250m)
16	Fresh water marshes	/		
17	Marshes	/		
18	Forest (mixed oak wood on marine levee)	VM	33	Natural dune forest
19	Natural floodplain forest	VM	35	Natural flood plain forest
		VM	37	Planted flood plain forest
20	Riparian forest	SP		shape along the rivers and channel
		CL		(tableland)
21	Forest on tableland	CL		
22	General forest in DDBR	VM	34	Planted dune forest
		VM	36	Natural flood plain forest with <i>Alnus glutinosa</i>
23	Bushes	VM	3	Coastal high dune (2-3m) vegetation
		VM	4	Coastal intermediate dune (1-2m) vegetation
		VM	5	Coastal low dune (0,5-1m) vegetation
		VM	8	Vegetation in depressions between dunes slightly salinised
		VM	11	Vegetation on flat marine / riverine deposits slightly salinised
		VM	13	Inland medium high dune vegetation
		VM	14	Inland low dune vegetation
		VM	38	Natural flood plain bushes
		VM	39	Natural flood plain bushes with dense <i>Tamarix ramosissima</i>
		VM	40	Planted flood plain bushes
24	Bushes on tableland	CL		
25	Tableland's foot	SP		
26	Marine levee	VM	2	Seashore vegetation

		VM	6	Vegetation in depressions between dunes strongly salinised
		VM	7	Vegetation in depressions between dunes medium salinised
		VM	12	Inland high dune vegetation
		VM	22	Mixed sedges vegetation on mineral soils
27	Meadows	VM	9	Vegetation on flat marine / riverine deposits strongly salinised
		VM	10	Vegetation on flat marine / riverine deposits medium salinised
		VM	16	Grassland on low river levee
		VM	17	Grassland on medium high river levee
		VM	18	Grassland on high river levee
		VM	23	Mixed Sedges vegetation on organic soils
28	Steppes	VM	15	Dry steppe vegetation on loess and rocks
29	Steppe on Tableland	CL		
30	Reed beds (with small willow formations)	VM	26	Mixed reed vegetation and bushes/trees on mineral soils
		VM	32	Reed vegetation and trees on open plaur
31	Reed beds (other)	VM	19	Pure reedmace vegetation on mineral soils
		VM	20	Mixed reedmace vegetation on mineral soils
		VM	21	Mixed reedmace vegetation on organic soils
		VM	24	Pure reed vegetation on mineral soils
		VM	25	Mixed reed vegetation on mineral soils
		VM	27	Mixed reed vegetation and bushes on organic soils
		VM	28	Reed vegetation on salinised soils
		VM	29	Reed vegetation on compact plaur
		VM	30	Reed vegetation and bushes on compact plaur
		VM	31	Reed vegetation on open plaur
		CL		
32	Floating aquatic vegetation (clairs de roselières)	CL		

VM	Vegetation map
CL	Classification
SP	Shape

Annexe 3 : Fiche de relevé terrain

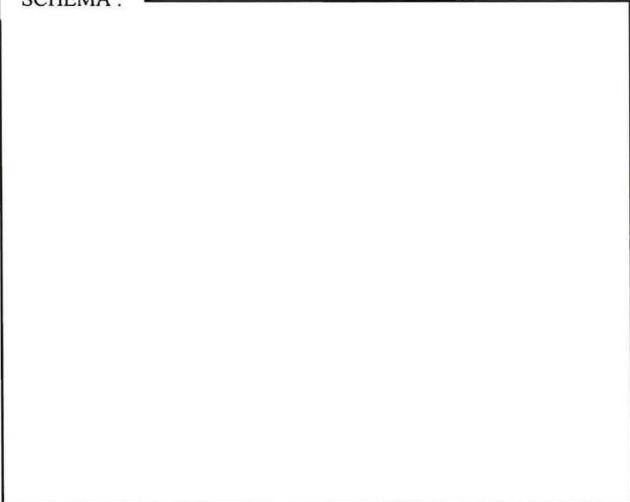
Vegetation sheet :

Site :

Zone :

Date :/ 06 / 2007

SCHEMA : _____



Remarks:

Geomorphology:

Talweg	Plain	Table land	Cliffs	riverbank	Valley	Crest	Hills

Soil occupation:

Not vegetative	Urban	Naked Soil	Sand	Rocks
Marine	Herbaceous Dune	Dune forest	Lagoon	Beach
Aquatic	Aquatic Herbaceous	River	Lake	
Marshes	Reed Beds	Scirpus	Rushes (junc)	
Meadows	Salt Marsh Veg	Grass (low/high)	Meadow	
Forest	Hard wood	Pine wood	Riparian Forest	Hedges Scrubland (guarigue)
Agricultural	Vineyard	Cereal	Riceland	Fruit Trees Fallow land

Annexe 4 : Equation du critère d'information d'akaike (AIC) et du critère d'information d'akaike corrige (AICc)

$$Prev = 0,19 + 0,69 \cdot URB + 0,51 \cdot EAU$$

Où L est la vraisemblance maximisée et k le nombre de paramètres du modèle

$$AICc = AIC + \frac{2 \cdot k \cdot (k+1)}{n - k - 1}$$

Où n est le nombre d'observations

Annexe 5 : Equation du coefficient kappa (K) pour l'évaluation d'une classification

$$K = \frac{\sum_{i=1}^n N \cdot X_{ii} - \sum_{i=1}^n X_{ip} \cdot X_{pi}}{N^2 - \sum_{i=1}^n X_{ip} \cdot X_{pi}}$$

Les variables de l'équation du coefficient Kappa sont explicitées dans le tableau suivant représentant une matrice de confusion :

	C11	Cii	Cnn	Total lignes
C1	X11	X 1i	X 1n	X1p
Ci	X i1	X ii	X in	Xip
Cn	X n1	X ni	X nn	Xnp
Total colonnes	Xp1	Xpi	Xpn	N

Annexe 6 : Glossaire épidémiologique

- Vecteur : organisme qui ne provoque pas lui-même une maladie mais qui disperse l'infection en transportant les agents pathogènes d'un hôte à l'autre. Des espèces de moustiques, par exemple, servent d'insecte vecteur pour le virus du Nil occidental qui provoque une maladie ; les insectes peuvent ingérer ce virus en se nourrissant d'un oiseau infecté et régurgiter dans un humain, en infectant donc celui-ci ou celle-ci.
- Hôte : organisme qui héberge un parasite, un partenaire mutuel ou un partenaire commensal, lui fournissant, en général, le gîte et le couvert.
- Zoonose : maladie animale, microbienne ou parasitaire, qui se transmet de l'animal à l'homme et vice versa (du grec zôon, « animal » et nosos, « maladie »).
- Arbovirus : type de virus ayant pour vecteur les arthropodes suceurs de sang : moustiques, tiques.
- Arboviroses : maladies virales dues à des arbovirus.
- Endémique : une maladie est dite endémique d'une région si elle y sévit de manière permanente.

Résumé :

Un grand nombre de maladies animales et humaines dépendent de facteurs environnementaux, pouvant être caractérisés par télédétection. Dans le cas d'une maladie multi-hôtes, comme la fièvre de West Nile, un grand nombre d'espèces (hôtes et vecteurs) peuvent être impliquées dans le processus de transmission.

Dans le cadre d'une étude sur la circulation du virus de West Nile en Roumanie, l'objectif du travail réalisé lors du stage est de caractériser la diversité des associations dans le temps et dans l'espace des espèces d'hôtes (oiseaux sauvages), de décrire la diversité de la communauté d'oiseaux et rechercher des facteurs de risques pour la transmission du Virus de West Nile.

Une carte de la végétation ainsi qu'une classification supervisée d'une image multispectrale à moyenne résolution spatiale ont permis de réaliser une carte des habitats oiseaux sur la région de Dobrogea en Roumanie. Couplée à une base de données oiseaux cette carte permet de localiser les différentes espèces potentiellement impliquées dans la circulation du virus de West Nile. Cet outil (carte des habitats couplé à la base de données oiseaux) a également permis de développer des cartes caractérisant la répartition et la diversité de la communauté oiseaux. Une analyse statistique des données environnementales par rapport à des mesures de séroprévalence ont permis de sélectionner les paramètres environnementaux pouvant avoir un effet amplificateur ou protecteur dans la circulation du virus.

Le travail réalisé ne constitue pas un modèle fiable permettant de prédire le risque de transmission, mais il est une première étape importante pour étudier les conditions environnementales favorables à la transmission du virus de West Nile. Il s'avère également être un outil indispensable pour l'unité ornithologique, et les études épidémiologiques à venir.

Mot clés : West Nile, delta du Danube, télédétection, SIG, carte d'habitats oiseaux, facteurs de risque

Summary:

A large number of animals and human diseases depend on environmental factors, which can be described by remote sensing. In the case of a multi-host disease, like West Nile Fever, a large number of species (host and vectors) is supposed to be involved in the transmission.

As part of a study on the West Nile circulation, the objectives of the realised work are to characterise the diversity of the host (wild birds) species associations in time and space; to describe the diversity of the bird community and to search for risk factors of the West Nile Virus transmission.

A vegetation map and a supervised classification of a multispectral image have been used to realise a bird habitat map on the Dobrogea area in Romania. Associated to a bird data base, this map allows to locate the different species potentially involved in the West Nile Virus circulation. This tool (habitat map linked to a bird database) has also permitted to develop maps that characterise the distribution and the diversity of the bird community. A statistic analysis of the environmental data compared to the seroprevalence measures has permitted to select the environmental parameters that can have a magnifying or protesting effect on the virus circulation.

The realised work does not consist to a reliable model for the transmission risk prediction, but it is the first important step to study the favourable environmental conditions to the West Nile Virus transmission. It also seems to be an essential tool for the ornithological unit, and for the next epidemiological studies.

Key words: West Nile Virus, Danube delta, remote sensing, GIS, bird habitat map, risk factors