



HAL
open science

Essais de modélisation de la présence de végétations naturelles à partir de variables environnementales et de données floristiques

Maelle Rambaud, Laurent L. Azuelos, Pierre Lafon, Gaël Causse, Hélène Bressaud

► To cite this version:

Maelle Rambaud, Laurent L. Azuelos, Pierre Lafon, Gaël Causse, Hélène Bressaud. Essais de modélisation de la présence de végétations naturelles à partir de variables environnementales et de données floristiques: Application à la région Île-de-France. [Rapport de recherche] CBNBP - MNHN, Délégation Ile-de-France, 61 rue Buffon - CP53 - 75005 PARIS cedex 05, France. 2013, pp.39. mnhn-02263824

HAL Id: mnhn-02263824

<https://mnhn.hal.science/mnhn-02263824>

Submitted on 5 Aug 2019

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

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



Conservatoire Botanique National



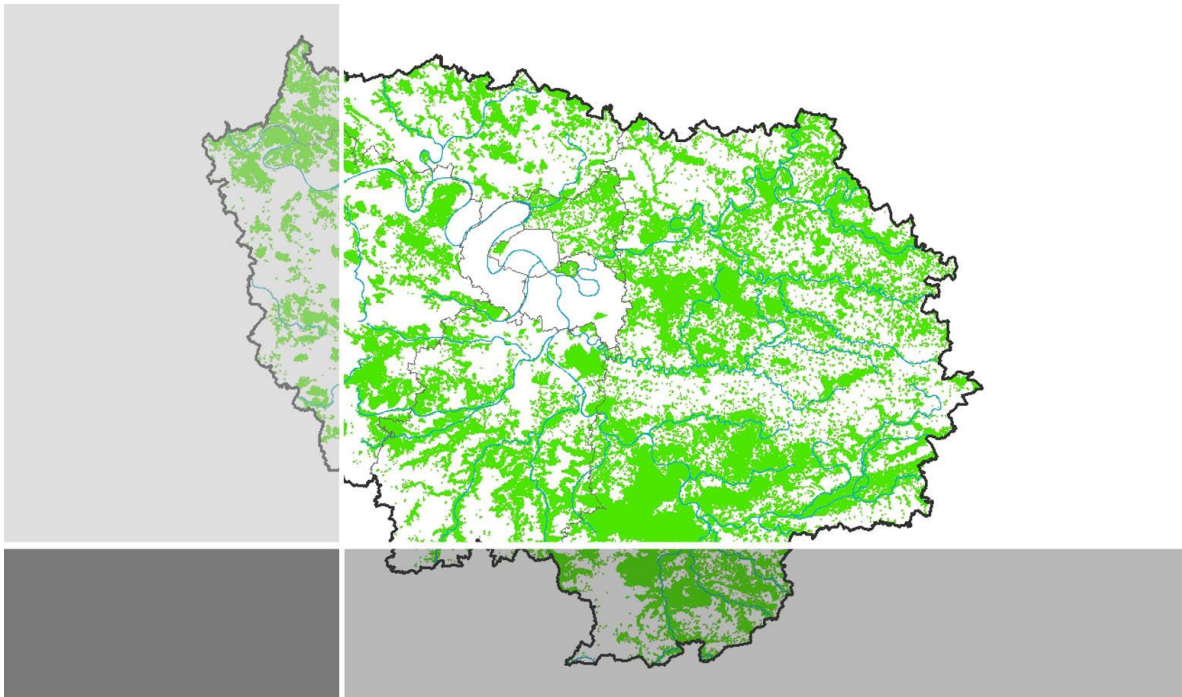
BASSIN PARISIEN

Conservatoire botanique national du Bassin parisien
Une structure au cœur du développement durable

Connaître
Comprendre
Conserver
Communiquer

Essais de modélisation de la présence de végétations naturelles à partir de variables environnementales et de données floristiques

Application à la région Île-de-France



Mars 2013

Conservatoire Botanique National



Conservatoire botanique national du Bassin parisien

USM 304 – Unité Inventaire et suivi de la biodiversité

Muséum national d'Histoire naturelle

61, rue Buffon - CP 53 - 75005 Paris – France

Tél. : 01 40 79 35 54 – cbnbp@mnhn.fr



Conservatoire botanique national du Bassin parisien
Une structure au cœur du développement durable

Connaître
Comprendre
Conserver
Communiquer



Conservatoire botanique national du Bassin parisien

Une structure au cœur du développement durable

Connaître

Comprendre

Conserver

Communiquer

Essais de modélisation de la présence de végétations naturelles à partir de variables environnementales et de données floristiques

Application à la région Île-de-France

Auteurs du rapport : Maëlle Rambaud, Laurent Azuelos, Pierre Lafon, Gaël Causse et Héléna Bressaud

CBNBP

Mars 2013.

Conservatoire botanique national du Bassin parisien

USM 304 – Unité Inventaire et suivi de la biodiversité

Muséum national d'Histoire naturelle

61, rue Buffon - CP 53 - 75005 Paris Cedex 05 – France

Tél. : 01 40 79 35 54 – cbnbp@mnhn.fr

Essais de modélisation de la présence de végétations naturelles à partir des variables environnementales et des données floristiques. Application à la région Île-de-France.

Maëlle Rambaud, Laurent Azuelos, Pierre Lafon, Gaël Causse & Héléna Bressaud - Mars 2013 - CBNBP

Essais de modélisation de la présence de végétations naturelles à partir de variables environnementales et de données floristiques

Application à la région Île-de-France

Ce document a été réalisé par le Conservatoire botanique national, du Bassin parisien, délégation Île-de-France, sous la responsabilité de

Frédéric Hendoux, directeur du Conservatoire
Conservatoire botanique national du Bassin Parisien
Muséum national d'Histoire naturelle
61 rue Buffon CP 53, 75005 Paris Cedex 05
Tel. : 01 40 79 35 54 – Fax : 01 40 79 35 53
E-mail : cbnbp@mnhn.fr

Étude menée pour le compte de :

Fédération des Conservatoires botaniques nationaux
10-18 rue Beaumarchais - BP 87
93511 Montreuil-sous-Bois Cedex

Rédaction et mise en page : Maëlle Rambaud, Laurent Azuelos, Pierre Lafon, Gaël Causse

Cartographie et saisie des données : Laurent Azuelos, Leslie Ferreira, Anaïse Bertran, Pierre Lafon, Ombeline Ménard, Aurélien Culat

Gestion et analyses des données : Maëlle Rambaud, Héléna Bressaud

Les partenaires de cette étude sont :

Conseil général de la Seine-et-Marne
Hôtel du département
77010 Melun cedex
Tél : 01.64.14.76.29



Direction régionale et interdépartementale de l'environnement
et de l'énergie - Driea
10, rue Crillon
75194 Paris cedex 04



Conseil régional d'Île-de-France
33, rue Barbet de Jouy
75007 Paris
Tél : 01.53.85.53.85



Ministère de l'écologie, du développement durable et de l'énergie
Grande Arche, Paroi Sud
92055 Paris-La-Défense Cedex



Crédit photo

Photo de couverture : © J. Monticolo, Cartographie des végétations de l'Île-de-France.

Essais de modélisation de la présence de végétations naturelles à partir des variables environnementales et des données floristiques. Application à la région Île-de-France.
Maëlle Rambaud, Laurent Azuelos, Pierre Lafon, Gaël Causse & Héléna Bressaud - Mars 2013 - CBNBP

Sommaire

Résumé	6
Introduction	8
1. Modélisation de végétations en Seine-et-Marne à partir de données de végétation et de l'Ecomos	10
1.1 – Méthodologie	10
1.1.1. Données disponibles	10
1.1.1.1. Données Végétation disponibles	10
1.1.1.2. Données environnementales disponibles	10
1.1.1.3. Fond blanc.....	11
1.1.2. Préparation et mise en forme des données.....	12
1.1.2.1. Croisement entre cartographies Végétation et Ecomos	12
1.1.2.2. Croisement des couches Végétation et Ecomos avec les petites régions naturelles	14
1.1.2.3. Croisement des couches Végétation et Ecomos avec les données environnementales	15
1.1.2.4. Sélection des données et des objets géographiques	17
1.1.3. Analyse des données et modèle statistique	18
1.1.4. Résumé du principe de modélisation	19
1.1.4.1. Préparation des données	19
1.1.4.2. Croisement des données avec les variables environnementales et analyses statistiques	19
1.1.4.3. Mise à jour des données attributaires.....	19
1.1.4.4. Estimation du taux de réussite associé à la modélisation.....	20
1.2 – Bilan provisoire - Perspectives	21
1.2.1. Bilan.....	21
1.2.2. Limites de la modélisation.....	21
1.2.2.1. Concordance géographique Végétation/Ecomos	21
1.2.2.2. Correspondances attributaires « Végétation-Ecomos »	22
1.2.3. Perspectives	22
2. Localisation de végétations dans la partie Ouest de l'Île-de-France à partir de données floristiques géolocalisées	24
2.1 – Cadre et objectif	24
2.2 – Méthodologie	24
2.2.1. Evolution par rapport à 2011	24
2.2.2. Préparation et mise en forme	25
2.2.2.1. Données utilisées	25
2.2.2.2. Méthodologie.....	26
2.2.2.3. Résultats de la requête	27
2.2.2.4. Prospection sur le terrain.....	28
2.3 – Bilan provisoire - Perspectives	29
2.3.1. Bilan.....	29
2.3.2. Limites de la méthode.....	30
2.3.3. Perspectives	31
Conclusion	32
Bibliographie	33
Annexe 1 : Types d'Ecomos (CLC6)	35
Annexe 2 : Script R	38

Résumé

Le programme Carhab contient dans sa feuille de route un volet modélisation, dont l'objectif est d'étudier l'apport des techniques de modélisation spatialisée de la végétation à la réalisation finale de la cartographie phytosociologique de France au 1/25 000^{ième}.

Dans le cadre de la mise en œuvre de la cartographie des végétations de la région Île-de-France, cofinancée par le Conseil Général de Seine-et-Marne, la Direction Régionale et Interdépartementale de l'Environnement et de l'Énergie d'Île-de-France (DRIEE-IF) et par le Conseil Régional d'Île-de-France, le Conservatoire botanique national du Bassin parisien (CBNBP) a mis en œuvre des méthodes d'automatisation et d'aide à la cartographie afin de tenir l'objectif d'une cartographie exhaustive, à l'horizon 2013, des milieux naturels et semi-naturels tels que définis par l'Institut d'Aménagement et d'Urbanisme (IAU 2005, Ecomos 2000). C'est ainsi que des outils informatiques ont été développés au sein du CBNBP pour renseigner les milieux ne pouvant faire l'objet d'inventaires de terrain dans le temps imparti à la Convention.

Deux méthodes ont été initiées dans des contextes différents et sont adaptées à la finalisation de la cartographie des végétations de la région.

Premièrement en Seine-et-Marne, le CBNBP réalise depuis 2006 la cartographie des végétations naturelles et semi-naturelles du département. Les cartes produites sont constituées de polygones cartographiés *in situ* ou *ex situ* et les végétations identifiées sont rattachées notamment à la nomenclature phytosociologique. Comme les délais sont courts, tout ce qui n'est pas cartographié fait l'objet d'une modélisation basée sur les polygones déjà cartographiés et sur des données environnementales. Cette modélisation surfacique est destinée à renseigner les polygones de l'Ecomos encore non cartographiés et donc sans affectation syntaxonomique.

La modélisation se décompose en sous-modèles : par petite région naturelle (croisement des couches des paysages de la Seine-et-Marne, des données géologiques, des données floristiques...), et par type d'Ecomos (ex : feuillus denses xéro à mésophile, code Clc6 31111).

Au sein d'un sous-modèle, les variables environnementales exploitées concernent la pente, l'altitude, l'exposition, un indice mesurant la densité du couvert végétal, les précipitations et la géologie (selon un gradient acide-calcaire). Une moyenne pour chacune de ces variables est calculée par polygone Végétation (dénomination utilisée pour caractériser les polygones cartographiés) et par polygone Ecomos. Les polygones Végétation et Ecomos avec leurs valeurs environnementales sont alors injectés dans une analyse en composante principale (ACP). Une distance entre polygones Végétation et polygones Ecomos est calculée dans un second temps à partir des combinaisons de variables environnementales issues de l'ACP. L'objectif est de déterminer au sein d'une petite région naturelle, pour chaque type d'Ecomos, quels polygones Végétation se rapprochent le plus de chaque polygone Ecomos du point de vue de ses conditions environnementales. Une fois les polygones Végétation identifiés, le ou les syntaxons majoritaires de ces polygones Végétation sont attribués au polygone Ecomos auquel ils ont été rattachés.

Un test sur un échantillon de données Végétation est finalement proposé pour mesurer l'efficacité de la modélisation et préconiser des améliorations en vue d'optimiser ses performances.

Essais de modélisation de la présence de végétations naturelles à partir des variables environnementales et des données floristiques. Application à la région Île-de-France.

Maëlle Rambaud, Laurent Azuelos, Pierre Lafon, Gaël Causse & Hélène Bressaud - Mars 2013 - CBNBP

Concernant le reste de la région Île-de-France (partie Ouest), les chargés d'étude ont élaboré une méthode applicable sur des types de végétations ponctuelles, souvent d'intérêt patrimonial (ex : végétations pionnières, fugaces, ourlets intraforestiers...), qui sont difficilement cartographiables.

L'objectif principal de la méthode est d'exploiter des données issues d'inventaires floristiques, beaucoup plus nombreuses que les données phytosociologiques, afin de localiser les végétations définies ci-dessus.

Pour ce faire, une liste de syntaxons représentatifs de ces végétations est retenue. Des taxons indicateurs de ces syntaxons sont ensuite sélectionnés. Dans un polygone cartographié et associé à une liste de taxons floristiques, si un nombre suffisant de taxons indicateurs est retrouvé, le syntaxon associé peut être attribué au polygone. D'autres critères servent aussi de filtres et peuvent aider dans l'attribution du syntaxon à un polygone : le code Corine biotopes, des taxons indicateurs négatifs, la taille du polygone, le nombre total de taxons, etc.

Comme pour la méthode précédente, des prospections de terrain sont réalisées en complément pour tester l'efficacité de la méthode et mesurer le taux d'erreur.

Avec les données obtenues dans ces deux méthodes et grâce à l'estimation du taux de réussite par méthode et par type de végétations, la cartographie d'Île-de-France des végétations naturelles et semi-naturelles peut être finalisée, tout en gardant un degré de fiabilité pour chaque polygone de végétations en fonction de l'origine de la donnée (inventaire, photo-interprétation, modélisation, localisation par liste floristique...). Ces deux méthodes pourraient tout à fait être déployées à l'échelle nationale dans le cadre de Carhab, car les données nécessaires au paramétrage et à la mise en œuvre des modèles (fond blanc physiognomique, données environnementales et données floristiques) sont ou seront bientôt disponibles sur l'ensemble du territoire métropolitain.

Introduction

Suite à un cofinancement sur la période 2008-2013 entre la Direction Régionale et Interdépartementale de l'Environnement et de l'Énergie d'Île-de-France (DRIEE-IF) et le Conseil Régional d'Île-de-France, superposé pour le département de la Seine-et-Marne à une convention avec le Conseil Général de Seine-et-Marne sur la période 2006-2012, le Conservatoire botanique national du Bassin parisien (CBNBP) s'est engagé à cartographier les végétations au 1/10 000^{ème}, sur la totalité des surfaces de l'Ecomos 2000 d'Île-de-France.

Cet Ecomos, réalisé par l'Institut d'Aménagement et d'Urbanisme (IAU 2005), correspond à une cartographie des grandes formations végétales naturelles et semi-naturelles de la région. Pour le CBNBP, il sert de fond blanc sur lequel il est nécessaire d'identifier le plus finement possible les végétations au sens phytosociologique. Ce fond blanc, qui constitue une pré-cartographie du territoire, est doté d'une typologie physionomico-écologique propre et en ce sens est comparable au fond blanc de Carhab dont la nomenclature est en voie de finalisation, et les tests de production en cours dans plusieurs départements (Cher, Loire, Isère).

Sur les 345 877 hectares de formations végétales couvertes par l'Ecomos que le CBNBP doit cartographier (173 118 dans le département de la Seine-et-Marne, 172 759 dans la partie Ouest de l'Île-de-France), 265 253 hectares (77,7%) le sont à l'heure actuelle (155 119 ha environ en Seine-et-Marne (89,6%), 110 404 ha dans la partie Ouest de l'Île-de-France (63,9%) sans compter les données de 2013). Cette cartographie provisoire fait suite à de nombreuses prospections de terrain et de longues phases de photo-interprétation. Les espaces restants constituant des végétations morcelées, associées aux bocages et aux cultures, souvent d'un moindre intérêt floristique ou phytosociologique, sont plus difficiles et chronophages à interpréter. Afin de pallier ce manque, une méthode de modélisation des végétations a été testée en Seine-et-Marne et sera étendue prochainement à l'ensemble de l'Île-de-France, et une méthode de localisation de végétations à partir de données floristiques a été testée dans la partie Ouest de la région.

Les deux méthodes doivent permettre d'attribuer une végétation potentielle à chaque polygone de l'Ecomos non cartographié, mais également d'ajouter des polygones de végétations non représentés sur l'Ecomos à l'échelle où il est restitué (de l'ordre du 1/20 000^{ème}). Pour la méthode de modélisation, différentes données sont nécessaires : la cartographie des végétations obtenue par prospection ou photo-interprétation et les données attributaires associées et stockées dans la base *Habitat* (données/polygones Habitat = données/polygones Végétation par la suite), des données de pente, d'altitude, d'exposition, de précipitation ainsi que des données mesurant la densité du couvert végétal et des données indiquant la géologie (selon un gradient acide-calcaire). Pour la méthode de localisation à partir de données floristiques, les données issues des relevés floristiques du CBNBP (données/polygones de la base de données *Flora*) sont utilisées et particulièrement intéressantes car très nombreuses.

Pour chaque méthode, des tests de comparaison sur des échantillons de jeux de données, ainsi que quelques jours de terrain sont envisagés, afin d'estimer le taux d'erreur et de valider ou d'infirmer la pertinence de chaque méthode.

Les apports de ces deux méthodes sont divers, notamment une application directe dans la convention CBNBP-DRIEE-Région, afin de compléter la carte des végétations d'Île-de-France, mais également dans le cadre du programme Carhab dont la première phase (2012-2014) de cadrage méthodologique a identifié la modélisation spatialisée des végétations comme un axe intéressant à développer, en complément à la cartographie de terrain, afin de 1/ intégrer des informations complémentaires au fond blanc, 2/ préparer et faciliter la réalisation terrain de la cartographie, 3/ détecter la présence de végétations ponctuelles couvrant des surfaces non cartographiables (feuille de route Carhab, partie A.2.3.).

1. Modélisation de végétations en Seine-et-Marne à partir de données de végétation et de l'Ecomos

1.1 – Méthodologie

1.1.1. Données disponibles

1.1.1.1. Données Végétation disponibles

Le CBNBP dispose en Seine-et-Marne d'une cartographie des végétations naturelles et semi-naturelles au 1/10 000^{ème} d'environ 165 745 ha, résultant de 6 ans de prospection et de photo-interprétation par 2 chargés d'études à temps plein sur le projet. Sur cette cartographie apparaissent des objets ponctuels inférieurs à 625 m² (0,0625 ha) et des objets surfaciques (sous forme de polygones) supérieurs à cette même surface. Un objet cartographié peut contenir une ou plusieurs végétations (dans ce deuxième cas le terme de « mosaïque » de végétations est employé). Le SIG MapInfo est utilisé pour construire la couche vectorielle regroupant l'ensemble de ces objets.

Sur le terrain, la récolte de données est effectuée *in situ* entre Avril et Septembre au moyen de relevés phytosociologiques ou de relevés floristiques (Azuelos & Renaud 2009). Certaines végétations sont également inventoriées à distance si elles sont difficiles d'accès mais suffisamment visibles. Au bureau, une partie importante de la cartographie est réalisée l'hiver par photo-interprétation. La digitalisation des objets géographiques est réalisée au 1/5 000^{ème} sur photos aériennes de 2003 croisées avec les cartes IGN-Scan 25 au 1/25 000^{ème}, les courbes altimétriques au 1/25 000^{ème} et les cartes géologiques au 1/50 000^{ème}.

Les informations syntaxonomiques sont les plus infra possibles (sous-association → classe en fonction de la capacité de détermination du syntaxon, souvent à l'alliance pour les données issues de la photo-interprétation) et les données attributaires nombreuses (conditions stationnelles, végétations identifiées selon la nomenclature phytosociologique, codes CORINE Biotopes et Natura 2000, état de conservation des végétations et relevés floristiques). Parmi celles-ci, l'information sur le syntaxon nous servira dans la méthode de modélisation des végétations.

1.1.1.2. Données environnementales disponibles

Les données environnementales disponibles ainsi que leur précision vont être primordiales pour accroître le succès de la modélisation. Différentes variables environnementales pertinentes ont été mentionnées par les chargés d'études grâce à leur expérience de terrain et leur connaissance de la

végétation. Malheureusement, nous ne disposons pas de toutes les couches pour toutes les variables concernées.

Au final, nous intégrons au modèle des variables de pente, d'altitude, d'exposition, de précipitation, des données sur l'indice de végétalisation représentant la densité du couvert végétal en Île-de-France, et des données basées sur la géologie exprimées sous forme d'un gradient acide-calcaire. L'idéal serait aussi d'ajouter des données sur la pédologie, l'humidité des sols, la climatologie locale... Il serait en particulier très intéressant d'ajouter certains indices, en cours d'utilisation et de test pour la mise en place du fond écologique de Carhab (Etlicher *et al.*, 2013), issus du modèle numérique de terrain (MNT) : indice d'humidité des sols, insolation, indice de position topographique.

Nous avons aussi, à une échelle plus grossière, une carte de petites régions naturelles relativement homogènes quant à leurs caractéristiques paysagères, géologiques et floristiques (Filoche *et al.* 2010). Les petites régions naturelles sont obtenues suite à des croisements d'information sur les paysages, la géologie et les affinités floristiques, ces affinités floristiques dérivant de calculs de similarité entre les relevés floristiques de la base de données *Flora* du CBNBP (cf. Filoche *et al.* 2010 pour l'explication de la méthode). Ainsi, dans la modélisation, nous travaillerons non plus à l'échelle du département dont les contours se basent sur des limites administratives uniquement, mais à l'échelle des petites régions naturelles dont les contours sont précisés par des caractéristiques écologiques et biologiques.

1.1.1.3. Fond blanc

L'Ecomos 2000 (IAU, 2005 ; 173 118 ha en Seine-et-Marne) sert de fond blanc sur lequel sont déterminés les polygones à prospector, photo-interpréter ou modéliser (cf. Figure 1). Cette cartographie simplifiée des milieux naturels d'Île-de-France est fondée sur l'interprétation combinée de photographies aériennes et d'images satellitaires sur l'ensemble des postes « naturels » du mode d'occupation du sol (MOS, IAU 2005) : bois ou forêts, coupes ou clairières en forêts, surfaces en herbe à caractère agricole, eaux fermées, surfaces en herbe non agricoles et espaces ruraux vacants. L'Ecomos ne traite donc pas des surfaces agricoles cultivées (terres labourées ou cultures spécialisées), des parcs et jardins ou des terrains de sports (http://www.iau-idf.fr/fileadmin/user_upload/SIG/presentation_ecomos.pdf). Sa nomenclature se décompose en niveaux de végétation emboîtés (6^{ème} niveau pour la plus grande précision, cf. Annexe 1 pour les types d'Ecomos les plus précis). Elle est issue et complémentaire de CORINE land cover (Heymann *et al.* 1994, Bossard *et al.* 2000).

Dans les missions du CBNBP, et notamment dans le cadre de la convention avec le CG77, les polygones de l'Ecomos ne sont pas suffisamment précis du point de vue géographique et typologique. Par conséquent, il nous incombe de les cartographier plus finement. Les phases de terrain, d'observations site par site, et les supports géographiques (notamment photo-aériennes, géologie, couche altimétrique) nous ont alors été indispensables pour affiner les postes et polygones de l'Ecomos. De plus, les végétations se devaient d'être identifiées d'un point de vue phytosociologique et non plus seulement en tant que simple formation végétale. En résumé, l'Ecomos a constitué une base très utile pour notre cartographie mais demandait à être affiné.

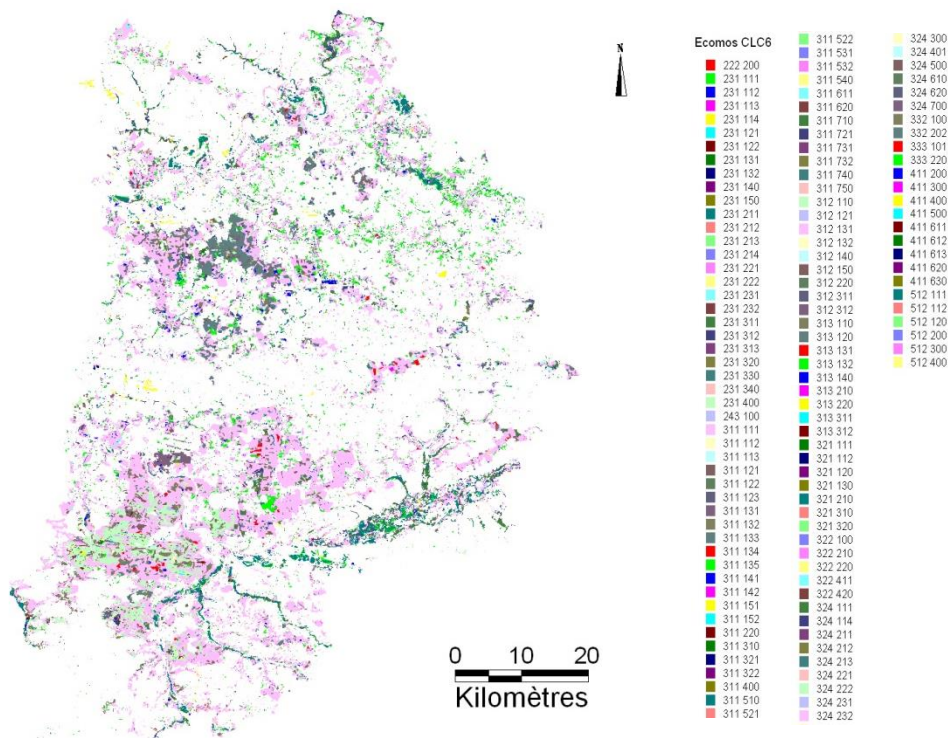


Figure 1 : Représentation cartographique de l'Ecomos en Seine-et-Marne, fond blanc à cartographier. 125 types d'Ecomos CLC6 sont présents sur le département. © CBNBP-MNHN, Rambaud 2012.

1.1.2. Préparation et mise en forme des données

1.1.2.1. Croisement entre cartographies Végétation et Ecomos

La première étape de la phase d'analyse consiste à croiser la cartographie des végétations naturelles et semi-naturelles du CBNBP obtenue par prospection et photo-interprétation (couche Végétation) avec la cartographie de l'Ecomos (cf. Figure 2). Au préalable, une sélection des polygones Ecomos supérieur à 2500 m² (0,25 ha) est réalisée, les autres polygones étant trop petits pour être suffisamment précis face à l'échelle de digitalisation de l'Ecomos. De même, l'Ecomos a été croisé avec le Mos afin d'exclure certains polygone Ecomos qui se superposent avec des polygones Mos très anthropiques (ex : prairies mésophiles chevauchant des parkings), ces biais étant dus à des erreurs de digitalisation essentiellement, ou plus rarement à une évolution des paysages.

L'objectif du croisement Végétation-Ecomos est de recenser puis de valider, pour chaque type d'Ecomos de précision maximale (6^{ème} niveau : CLC6), les syntaxons susceptibles d'y être observés et le nombre d'occurrence de ces syntaxons sur chaque type d'Ecomos. Différents syntaxons peuvent alors être associés au même type d'Ecomos puisque la classification des végétations selon la phytosociologie et selon l'Ecomos diffère.

Le pourcentage de recouvrement du polygone Végétation sur le polygone Ecomos est également renseigné afin de voir si le syntaxon est plus ou moins bien représenté sur le type d'Ecomos concerné par le polygone.

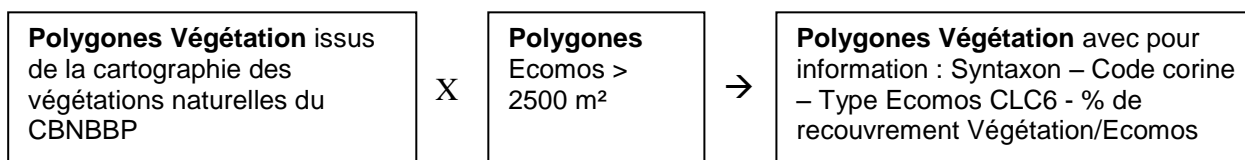


Figure 2 : Schéma du croisement entre cartographie des végétations naturelles et Ecomos.

Seules les associations polygone Végétation-polygone Ecomos dont le pourcentage de recouvrement du polygone Végétation sur l'Ecomos est supérieur à 1% (seuil lié à des erreurs de cartographie, ou d'imprécisions de l'Ecomos notamment) sont conservées. Il est alors intéressant de hiérarchiser la correspondance « Végétation-Ecomos » via les pourcentages de recouvrement (cf. Figure 3).

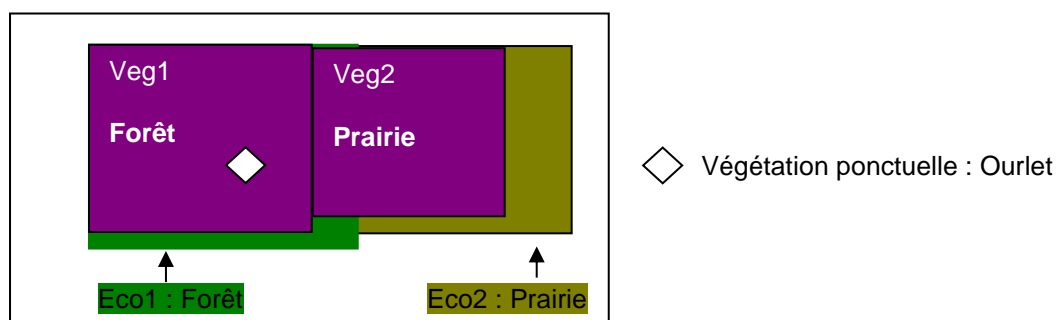


Figure 3 : Schéma du recouvrement des polygones Végétation (Veg1, Veg2) sur les polygones Ecomos (Eco1, Eco2).

A partir de ce croisement, les chargés d'études ont pour mission de valider ou d'infirmer les correspondances Syntaxons-Ecomos obtenues. En effet, un certain nombre de correspondances sont délicates ou douteuses, notamment car les polygones Ecomos (digitalisation au 1/10 000^{ème} avec peu d'échantillonnage de terrain et une typologie très simplifiée) sont moins précis que les polygones Végétation (digitalisation au 1/5000^{ème} avec beaucoup d'échantillonnage de terrain et une typologie fine), il peut donc y avoir des erreurs dans l'interprétation des types d'Ecomos. De manière synthétique, les correspondances problématiques sont :

- les correspondances valides mais à faible occurrence (moins de 10 % des polygones Ecomos occupés par un même syntaxon) qui risquent d'être non pertinentes dans la modélisation. Par exemple, le type Ecomos « des boisements de jeunes feuillus xéro à mésophiles » ne recoupe le *Salici cinerea* – *Rhamnion cathartica* que dans 4 % des cas.
- les correspondances impossibles dues à l'hétérogénéité d'échelle de digitalisation entre les couches Végétation et Ecomos. Par exemple, dans la Figure 3, l'analyse de la classe Ecomos « Forêt mésophile dense » (=Eco1) met en évidence :
 - des correspondances valides :
 - Forêt : *Carici flacca* - *Fagetum sylvatica* Thill 1964
 - Fourrés : *Tamo communis* – *Viburnetum lantanae* Géhu, Delelis & Frileux in Delelis ex J.-M. Royer, Felzines, Missot & Thévenin 2006
 - des correspondances impossibles :
 - Prairie : *Cynosurion cristati* Tüxen 1947.

Ce problème de superposition des polygones peut être résolu en travaillant avec des données ponctuelles Végétation (centroïdes) - 1 point Végétation = 1 information syntaxon et 1 type Ecomos.

- les correspondances non extrapolables dues aux végétations ponctuelles non identifiées par l'Ecomos. Il s'agit par exemple des ourlets calcicoles à acidiphiles des *Trifolio medii* - *Geranietea sanguinei* T. Müll. 1962, des ourlets acidiphiles des *Melampyro pratensis* - *Holcetea mollis* H. Passarge 1994, des végétations sur dalle rocheuse de *Alyso alyssoidis* - *Sedion albi* Oberd. & T. Müll. in T. Müll. 1961, des végétations amphibies des *Isoeto durieui* - *Juncetea bufonii* Braun-Blanq. & Tüxen ex V. West., J. Dijk & Paschier 1946 et des *Littorelletea uniflorae* Braun-Blanq. & Tüxen ex V. West., J. Dijk & Paschier 1946.

Dans l'exemple de la Figure 3, les ourlets sont confondus dans les classes forestières de l'Ecomos. La modélisation ne doit donc concerner que les végétations surfaciques.

- les correspondances issues des mosaïques de végétations. Dans l'exemple de la Figure 3, nous pourrions avoir une mosaïque de forêts, de fourrés et de clairières forestières pour la même classe d'Ecomos. En fonction des végétations modélisées, le lecteur devra donc supposer des potentialités de complexes de végétations dynamiques et/ou topographiques.
- les classes Ecomos fourre-tout (ex : végétation arbustive en mutation, végétation clairsemée). Chacune de ces classes présente une variabilité de végétations observées *in situ* trop importante et impossible à modéliser. Ces polygones correspondent à des milieux en transition mal définis et seront exclus de la modélisation.

1.1.2.2. Croisement des couches Végétation et Ecomos avec les petites régions naturelles

Les petites régions naturelles du département de la Seine-et-Marne (et de toute l'Île-de-France d'ailleurs) ont été définies en interne au sein du CBNBP. Le travail s'est appuyé sur de la bibliographie (régions phyto-écologiques, Dupias & Rey, 1985 ; petites régions agricoles, D.R.A.F., 1992 ; cartes des Paysages de la Seine-et-Marne, CG77, 2009) et sur le croisement de diverses couches raster portant sur les conditions géologiques, topographiques, hydrographiques, climatiques et parfois anthropiques. Ces variables environnementales ont toutes plus ou moins d'influence sur la nature de la végétation. Une contrainte a toutefois été imposée : les petites régions naturelles devaient se baser sur les limites communales (Filoche *et al.* 2010).

Des changements ont été apportés à la carte de petites régions naturelles obtenue après bibliographie et croisement géographique grâce à une analyse basée sur des affinités floristiques. Cette analyse repose sur la similarité entre communes en termes de composition floristique. Plus les communes partagent des espèces en commun, plus elles se ressemblent du point de vue des conditions écologiques susceptibles d'accueillir ces espèces. Des classes de communes avec des compositions floristiques proches ont ainsi été établies. Ces classes ont permis de préciser l'appartenance des communes aux petites régions naturelles en fonction des conditions écologiques mises en avant par les préférences des espèces végétales, et non plus seulement en fonction des variables environnementales (cf. Figure 4).

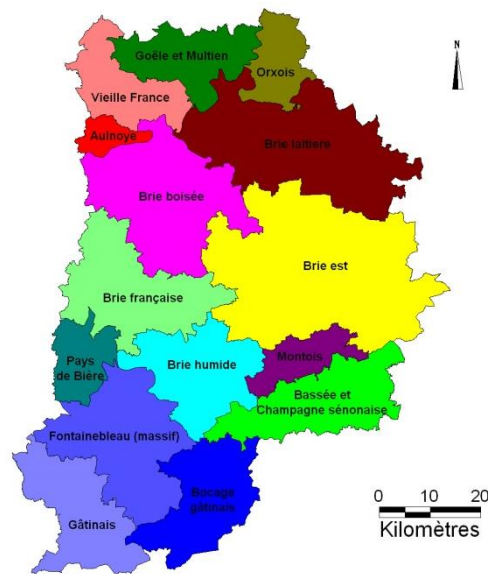


Figure 4 : Carte des petites régions naturelles de la Seine-et-Marne. © CBNBP-MNH, Rambaud 2012.

Le croisement des couches Végétation et Ecomos avec la couche des petites régions naturelles nous permet ensuite d'attribuer une petite région naturelle à chaque objet géographique Végétation ou Ecomos (cf. Figure 5). Si l'objet se situe à cheval sur plusieurs régions, celle contenant le centroïde de l'objet est sélectionnée.

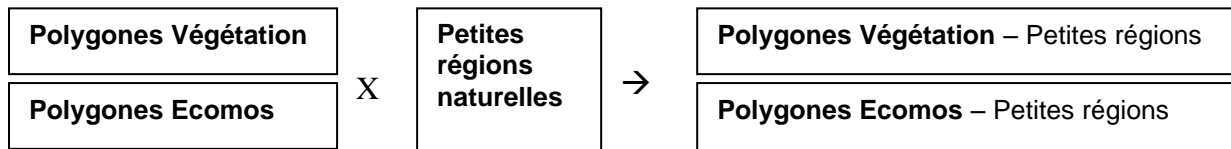


Figure 5 : Schéma des croisements des cartographies Végétation et Ecomos avec les petites régions naturelles de Seine-et-Marne.

L'information sur les petites régions naturelles est assez complexe. Elle servira à travailler à une échelle plus fine que celle du département, et ainsi à améliorer la modélisation en se basant sur des limites identifiées en partie sur des données écologiques et biologiques, et plus seulement administratives. En travaillant ainsi par petite région naturelle, nous sélectionnons un pool de végétations spécifiques correspondant à des complexes géomorphologiques et écologiques de même nature. Le choix d'affectation des polygones Ecomos vides sera ainsi plus restreint.

1.1.2.3. Croisement des couches Végétation et Ecomos avec les données environnementales

Les données environnementales dont nous disposons concernent la pente, l'altitude, l'exposition, les précipitations, un indice de végétalisation et la géologie.

Pente, altitude, et exposition sont renseignées dans la BD_Alti à l'échelle du pixel de 25 m (Institut Géographique National (IGN-F), BD ALTI® Base de Données altimétriques, Modèle Numérique de terrain (MNT), Imagerie Aster, <http://www.ign.fr/>). Les valeurs de pente et d'altitude sont numériques

(degrés et mètres). L'exposition est par contre une variable qualitative qui a été transformée en angle puis recodée en classe (cf. Figure 6).

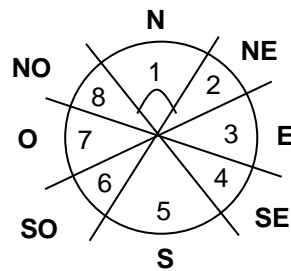


Figure 6 : Schéma expliquant comment les directions cardinales issues de l'exposition ont été recodées en classe.

Les précipitations sont fournies par Météo France et représentées sur une grille dont les points sont espacés de 1 km. Les données par point correspondent aux précipitations annuelles cumulées entre 1971 et 2011. A partir de la grille, un maillage avec les points comme centroïdes est construit (outil GridMaker de MapInfo), afin que tout l'espace ait une valeur de précipitation et pas seulement les points de la grille. L'indice de végétalisation (NDVI, IAU, Imagerie Spot, 2000, www.iau-idf.fr/fileadmin/Etudes/etude_799/doc/memento_1_Indice-vegetation.pdf) est quant à lui issu d'un raster dont la taille du pixel est de 15 x 15 m. Il reflète l'activité chlorophyllienne et la biomasse végétale, et il est codé en classe allant de 0 à 12. Plus sa valeur est élevée, plus la densité du couvert végétal est faible. La géologie pour finir a été renseignée pour chaque polygone par les chargés d'études selon un gradient acide-calcaire allant de 1 à 15 (ex : 2 = Sables acides *versus* 11 = Marnes).

La couche Végétation est alors croisée avec la BD-Alti dans un premier temps, avec la couche des précipitations dans un deuxième temps et avec l'indice de végétalisation dans un troisième temps. Le même type de croisement est ensuite effectué entre l'Ecomos et ces trois couches (cf. Figure 7). Comme les polygones Végétation et Ecomos comportent plusieurs pixels, une valeur moyenne par variable est calculée pour chaque polygone Végétation et Ecomos grâce à l'outil de statistiques zonales dans Spatial Analyst de ArcGis (cf. Figure 8).

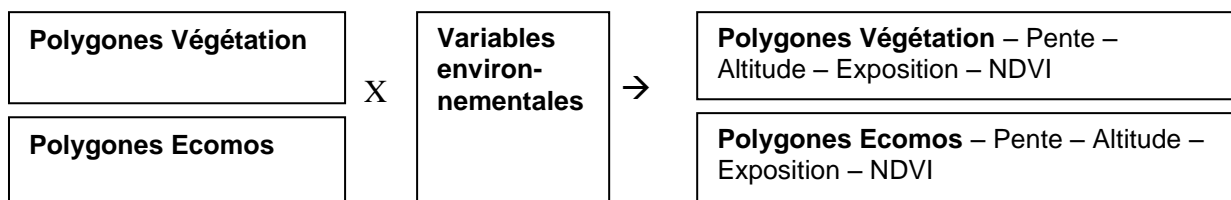


Figure 7 : Schéma des croisements des Végétations et de l'Ecomos avec les variables environnementales.

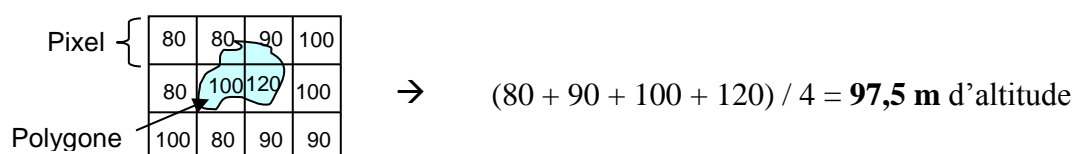


Figure 8 : Schéma montrant comment est calculée l'altitude moyenne d'un polygone sur une grille raster.

1.1.2.4. Sélection des données et des objets géographiques

Données Végétation

Tous les objets cartographiés par les chargés d'études ne sont pas utilisés dans la modélisation. Seules les prospections sur le terrain (*in situ*) sont prises en compte afin d'optimiser la validité des données analysées. La photo-interprétation est donc écartée. De même, seuls les polygones Végétation sont sélectionnés. Les objets ponctuels ne sont pas considérés car ils ne sont pas pris en compte par l'Ecomos (cf. 1.2.1). De plus, seules les correspondances syntaxon-Ecomos non douteuses sont gardées, ainsi que les polygones dont le pourcentage de recouvrement de la végétation sur l'Ecomos est supérieur à 1% (seuil lié à des erreurs de cartographie, ou d'imprécisions de l'Ecomos notamment). Le seuil des 1% peut éventuellement être relevé. Ce pool de données est considéré comme le pool de données valides.

A propos des végétations en mosaïque, pour ce premier essai de modélisation, le modèle nous permet de garder uniquement les syntaxons avec la plus grande surface au sein de la mosaïque. Nous travaillons donc avec un seul syntaxon par polygone. Par la suite, tous les syntaxons de la mosaïque devront être pris en compte. De même, les polygones où le syntaxon n'est pas renseigné, comme le cas peut se produire dans les cultures, sont pour l'instant écartés.

Polygones Ecomos

Les polygones Ecomos nous intéressant sont ceux qui ne comportent pas d'objets géographiques avec des données attributaires valides dans la couche Végétation. Ces polygones sont obtenus en croisant la couche Végétation valide avec la couche Ecomos, et en prenant l'inverse de ce croisement, c'est à dire en supprimant tous les polygones Ecomos qui contiennent des données Végétation valides. Nous récupérons ainsi une couche Ecomos_sans_Polygone_Végétation.

Evaluation de la modélisation

Un test est réalisé en retirant 30% des données Végétation valides. Cet échantillon est intégré dans la modélisation au même titre que les polygones Ecomos_sans_Polygone_Végétation. Ces polygones tests sont modélisés afin de servir de témoin pour évaluer l'efficacité de la modélisation. Pour cela, il suffit de comparer le syntaxon qui leur est attribué par modélisation et le syntaxon renseigné sur le terrain par les chargés d'études. En calculant le pourcentage d'adéquation entre la modélisation et les données des chargés d'études, nous obtenons une estimation du taux de réussite de la modélisation. Cette étape permet également de faire ressortir les syntaxons qui sont les mieux et les moins bien modélisés, et à partir de quel niveau syntaxonomique la modélisation entraîne trop de biais.

Remarque : Une vérification sur le terrain serait également très pertinente pour confirmer ou infirmer les syntaxons obtenus par modélisation pour quelques polygones Ecomos_sans_Polygone_Végétation.

1.1.3. Analyse des données et modèle statistique

Deux jeux de données entrent dans la partie statistique de la modélisation :

- le jeu de données Végétation valides avec les informations portant sur l'identifiant station/polygone, l'identifiant Végétation, le syntaxon, le type d'ecomos et son pourcentage de recouvrement, la petite région naturelle, les valeurs des variables environnementales ;
- le jeu de données associé aux polygones Ecomos_sans_polygone_Végétation avec les informations portant sur l'identifiant Ecomos, le type d'Ecomos, la petite région naturelle, les valeurs des variables environnementales.

Ces deux jeux de données sont intégrés dans le logiciel de statistique R (R Development Core Team 2008, cf. script Annexe 2) sous forme de tableaux avec les identifiants en ligne et les variables en colonne. Une ligne correspond donc à un identifiant Végétation unique ou un identifiant Ecomos unique avec toutes leurs informations associées sur la même ligne.

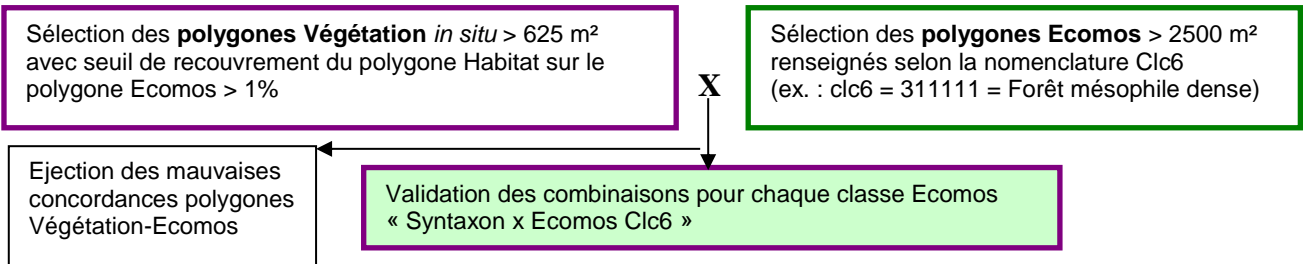
La modélisation se décompose en sous-modèles puisque nous travaillons par petite région naturelle et par type d'Ecomos (CLC6) au sein du département de la Seine-et-Marne. Nous avons donc un sous-modèle pour chaque combinaison « petite région-EcomosCLC6 ».

La première étape de chaque sous-modèle consiste en la réalisation d'une analyse en composante principale (ACP) sur les données Végétation pour obtenir une combinaison de variables par identifiant Végétation (plus facile à traiter que des variables prises séparément). Les données Ecomos sont ajoutées dans un deuxième temps en tant que variables supplémentaires. Les objets liés aux identifiants Végétation et Ecomos sont ainsi projetés dans un nouvel espace et positionnés en fonction de leur combinaison de variables. Une distance entre chaque identifiant Végétation et Ecomos peut alors être calculée. Nous récupérons ensuite, pour chaque identifiant Ecomos, l'identifiant Végétation (et les données associées) qui se rapproche le plus (distance la plus courte) en termes de combinaison de variables. À terme, si l'identifiant Végétation sélectionné se situe dans une station/polygone comportant plusieurs syntaxons, toute la mosaïque devra être attribuée à l'identifiant Ecomos. L'identifiant Ecomos pourra ainsi contenir plusieurs syntaxons.

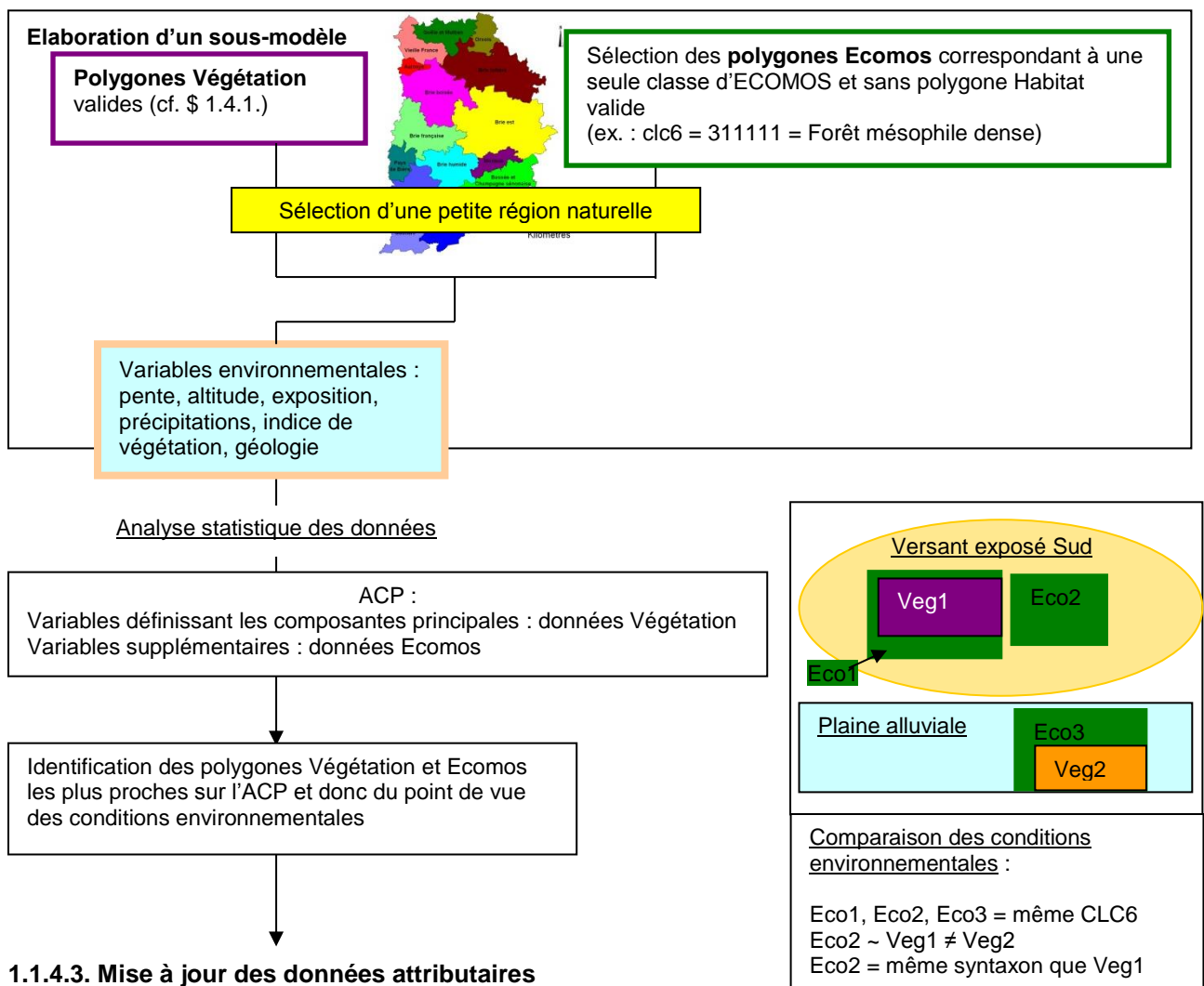
Une comparaison finale au sein de l'échantillon entre les syntaxons obtenus par modélisation et ceux renseignés par les chargés d'études permet d'estimer le taux d'erreur associé à la modélisation (cf. § 1.2.4.3).

1.1.4. Résumé du principe de modélisation

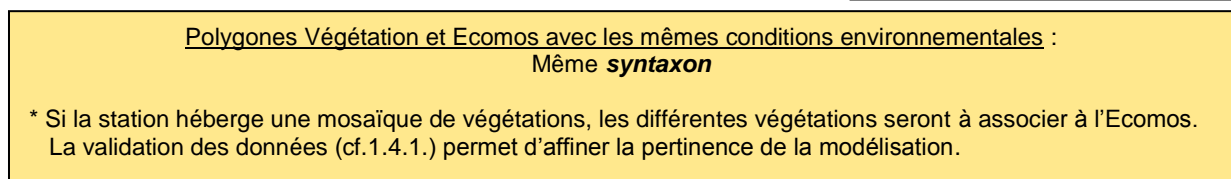
1.1.4.1. Préparation des données



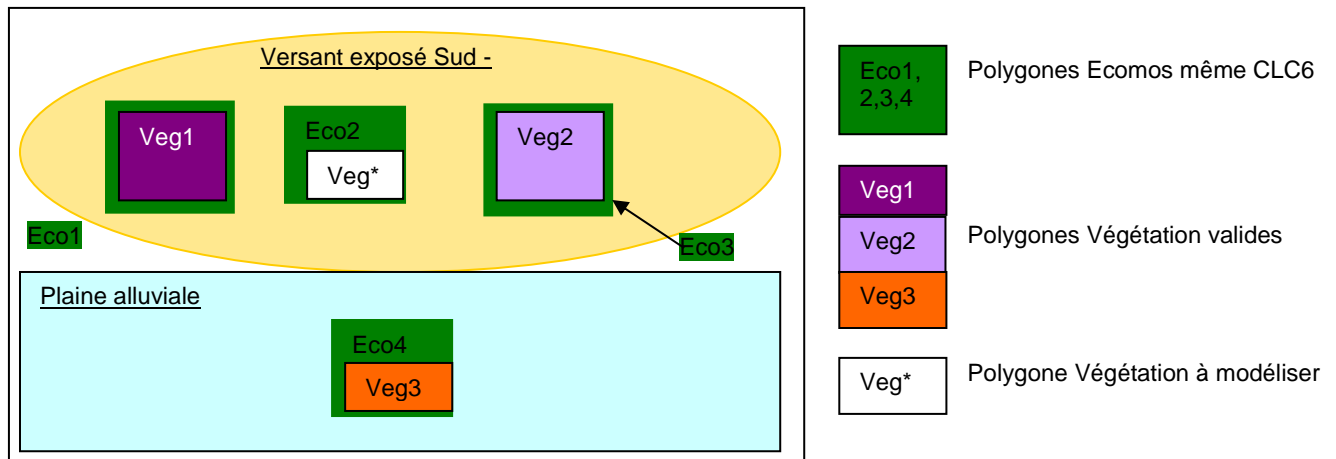
1.1.4.2. Croisement des données avec les variables environnementales et analyses statistiques



1.1.4.3. Mise à jour des données attributaires



1.1.4.4. Estimation du taux de réussite associé à la modélisation



1. Mise en évidence de variables environnementales proches entre Veg1~ Veg*.
2. Attribution du syntaxon de Veg1 à Veg*.
3. Comparaison du syntaxon de Veg* obtenu par modélisation et du syntaxon attribué sur le terrain par les chargés d'études.
4. Evaluation du taux de réussite associé à la modélisation (Veg* modélisée = Veg* chargés d'études ?).

1.2 – Bilan provisoire - Perspectives

1.2.1. Bilan

Dans le temps imparti à la cartographie des végétations, 17 999 ha (10,4%) de la surface de l'Ecomos n'ont pu être cartographiés par prospection ou par photo-interprétation. La modélisation nous a permis de modéliser 9805 ha, c'est-à-dire 5,7% de l'Ecomos supplémentaires. Les 8194 ha restants (4,7%) n'ont pas été intégrés à la démarche de modélisation, soit parce que les polygones Ecomos étaient trop anthropiques (ex : certaines cultures, les bords de routes, les polygones dont la digitalisation après croisement avec le Mos restait douteuse...), soit parce que l'Ecomos n'était pas modélisable (notamment dans le cas des très petits polygones).

Le test sur un échantillon de 30% des données Végétation valides a montré que le taux de réussite moyen de la modélisation est de 76% lorsque les syntaxons sont renseignés ou remontés au niveau de la classe phytosociologique (sauf pour les forêts pour lesquelles il était plus pertinent que l'information soit décrite à l'alliance) et de 50% pour une modélisation majoritairement à l'alliance phytosociologique (sauf pour les forêts décrites à l'association).

Les végétations présentant les meilleurs résultats sont bien entendu celles qui ont été suffisamment échantillonnées. Il s'agit des végétations spécifiques à une petite région naturelle (ex : les forêts alluviales de la Bassée dont le taux de réussite est monté à 88%), les végétations spécifiques à un groupe de classes d'Ecomos (ex : clairières forestières, landes sèches, roselières = taux de 67%) ou les végétations spécifiques à des conditions environnementales marquées comme les pentes fortes exposées au sud (ex : pelouses calcicoles sèches = taux de 70% et fourrés calcicoles secs = taux de 51%). La combinaison de ces trois variables a également permis d'avoir des taux de réussite satisfaisants pour les forêts et fourrés humides (70%), les friches (56%) et les prairies (63%). L'amélioration du taux de réussite de la modélisation des forêts non humides (45%) nécessiterait d'utiliser d'autres variables environnementales comme la pédologie. Pour les prairies, le registre parcellaire graphique (système d'information géographique permettant l'identification des parcelles agricoles) serait également intéressant afin de distinguer les prairies pâturées des prairies fauchées.

1.2.2. Limites de la modélisation

1.2.2.1. Concordance géographique Végétation/Ecomos

Le croisement des couches Végétation et Ecomos a mis en évidence une mauvaise concordance des deux couches géographiques. L'hétérogénéité de l'échelle de saisie induit ces erreurs de correspondance. Par conséquent pour une classe d'Ecomos de forêt, nous pouvons nous retrouver avec un syntaxon correspondant à une prairie.

Dans le but de contourner ce problème, il faudrait revoir certains paramètres de l'analyse :

1. manipuler les centroïdes des polygones Végétation à la place des polygones surfaciques ou ;
2. augmenter le seuil de recouvrement du polygone Végétation sur le polygone Ecomos (actuellement à 1%) pour que le polygone Végétation puisse être conservé dans l'analyse.

1.2.2.2. Correspondances attributaires « Végétation-Ecomos »

Le croisement entre les couches Végétation et Ecomos a également mis en évidence une grande diversité de correspondance due à la précision syntaxonomique du terrain. Pour une classe d'Ecomos, le nombre de possibilités de syntaxons étant trop grand, il s'avère nécessaire de procéder à une généralisation du niveau syntaxonomique.

La qualité des données étant moindre par modélisation, les syntaxons sont alors à remonter au niveau de l'alliance phytosociologique, sauf pour les forêts où une précision à l'association est nécessaire. Dans le cas où, pour une classe d'Ecomos avec conditions environnementales spécifiques, il existe encore une trop forte variabilité des alliances phytosociologiques, les syntaxons sont alors à remonter au niveau de la classe phytosociologique. C'est le cas par exemple des prairies mésophiles rattachées aux *Arrhenatheretea elatioris* Braun-Blanq. ex Braun-Blanq. et al. 1952 pour lesquelles il est difficile de distinguer les pâtures (ordre des *Trifolio repentis* – *Phleetalia pratensis* H. Passarge 1969) de prairies de fauche (ordre des *Arrhenatheretalia elatioris* Tüxen 1931). Il en sera de même pour les herbiers enracinés des eaux douces rattachés aux *Potametea pectinati* Klika in Klika & V. Novák 1941 pour lesquels une interprétation *in situ* (structure de la végétation et niveau trophique) est nécessaire afin de descendre au niveau de l'alliance ou de l'association.

Comme vu dans les paragraphes précédents, les végétations ponctuelles (ourlets, dalles rocheuses, végétations amphibies, etc.) n'étant pas identifiés par l'Ecomos, la modélisation ne concernera que les végétations surfaciques. D'autre part, les polygones correspondant à des milieux en transition mal définis (classes Ecomos « fourre tout ») sont également à exclure de la modélisation.

Pour finir, en fonction des végétations modélisées, le lecteur devra supposer des potentialités de complexes de végétations dynamiques et/ou topographiques (ex. : clairières forestières et ourlets au sein de forêts denses non identifiés par Ecomos).

1.2.3. Perspectives

Plusieurs pistes d'amélioration de la méthode sont envisageables.

Pour la partie cartographique, la récupération d'autres couches, telle que la pédologie, serait une énorme valeur ajoutée pour affiner le modèle via l'apport de variables environnementales pertinentes dans la détection des végétations naturelles. De même, d'autres indices dérivés du MNT (indice d'humidité des sols, indice topographique, insolation, ombrage...) pourraient être utilisés.

Autre point envisagé sur l'aspect cartographique, une solution pour contourner le problème de la non concordance des polygones Végétation et Ecomos. Il est peut-être utile de travailler avec des

données syntaxonomiques ponctuelles (centroïdes des polygones) plutôt que surfaciques (polygones), permettant ainsi de s'affranchir du problème de concordance cartographique entre les couches Ecomos et Végétation (cf. 2.2.1.).

Pour la partie statistique, il est envisageable de prendre les 5-10 premiers identifiants Végétation avec la distance la plus courte (et non plus seulement un seul identifiant) et de regarder, par identifiant Ecomos, le syntaxon qui ressort le plus souvent dans les 5-10 premiers identifiants Végétation. S'il y en a plusieurs, tous peuvent être associés à l'identifiant Ecomos et mettre ainsi en avant une incertitude sur la modélisation.

Il est également envisageable de faire une classification ascendante hiérarchique (CAH) après l'ACP à l'échelle de la Seine-et-Marne ou par petite région naturelle, et par type d'Ecomos (CLC6), pour classer les identifiants Végétation et Ecomos dans des groupes et voir dans ces groupes quels syntaxons ressortent le plus souvent. Il sera ainsi possible d'établir le % de présence de chaque syntaxon par groupe et donc par identifiant Ecomos appartenant à ce groupe. Ce test compléterait celui du paragraphe 1.2.1. en ajoutant l'influence des variables environnementales au simple croisement Végétation-Ecomos.

Il pourrait aussi être intéressant de coupler des données quantitatives et qualitatives à l'aide des données de Flora. Une réflexion est donc à prévoir quant à l'ajout d'une distance biologique basée sur les relevés floristiques de *Flora*. Le problème majeur est alors de coupler des données quantitatives et qualitatives. Deux options sont proposées pour se servir des relevés floristiques :

- à l'entrée de la modélisation : détection indirecte des alliances là où il y a des données de *Flora* disponibles (cf. méthode présentée en partie 2) ;
- en filtre de sortie de modélisation, permettant ainsi de rejeter l'affectation des polygones Ecomos dont les données floristiques ne correspondraient pas au syntaxon extrapolé via les données environnementales.

Pour le rendu final, sur la carte de la Seine-et-Marne qui contient les polygones Végétation renseignés par les chargés d'études et les polygones Ecomos renseignés par modélisation, un indice de fiabilité sera attribué à chaque polygone pour informer l'utilisateur de l'origine et de la qualité de la donnée.

2. Localisation de végétations dans la partie Ouest de l'Île-de-France à partir de données floristiques géolocalisées

2.1 – Cadre et objectif

Depuis 2008, le CBNBP élabore la cartographie des végétations naturelles et semi-naturelles d'Île-de-France sur les départements des Hauts-de-Seine, du Val-de-Marne, de Paris, des Yvelines, de l'Essonne et du Val-d'Oise. Cette cartographie, établie au 1/10 000^{ème}, est réalisée en deux étapes. La première consiste en un échantillonnage sur le terrain des différentes végétations d'un secteur. Les données peuvent être associées à des relevés floristiques ou phytosociologiques. La deuxième étape du programme de cartographie consiste en l'extrapolation par photo-interprétation des données acquises sur le terrain sur des secteurs de mêmes conditions écologiques (géologie, exposition...). Malgré tout, de nombreux types de végétations ponctuelles, notamment d'intérêt patrimonial (végétations pionnières, furtives, ourlets intraforestiers...), restent encore difficiles à cartographier. L'objectif principal de la méthode de localisation à partir de données floristiques est alors de tester la possibilité d'exploiter les données floristiques, beaucoup plus nombreuses que les données phytosociologiques proprement dites, afin de localiser les végétations restantes et de compléter l'information cartographique disponible.

Cette méthode a été testée pour la première fois lors de la campagne de terrain de 2011 et a donné des résultats intéressants et prometteurs (Lafon, 2012). Elle est par conséquent de nouveau utilisée en 2012 après avoir été améliorée grâce aux résultats de l'année 2011.

2.2 – Méthodologie

2.2.1. Evolution par rapport à 2011

L'année 2011 ne peut être considérée que comme une année test. En effet, elle a permis de montrer que l'utilisation des nombreux relevés floristiques contenus dans la base Flora du CBNBP, pouvait servir de base pour une localisation cartographique des végétations, mais a également mis en évidence un certain nombre de carences dans la démarche.

La méthode utilisée en 2012 a donc été améliorée avec les enseignements qu'avait apportés le test de 2011.

Ainsi, l'étude a été élargie à davantage de végétations, qui concernent tout les grands types de milieux. De plus, le niveau de recherche a été abaissé à l'association et non plus à l'alliance, comme c'était le cas en 2011. Cela n'a été réalisé que pour les associations où un résultat semblait possible à dire d'expert, les végétations à flore trop ubiquiste étant éliminées. Enfin, la méthode a également été étendue aux végétations moins patrimoniales.

2.2.2. Préparation et mise en forme

2.2.2.1. Données utilisées

Cette méthode se base sur l'important jeu de données floristiques que contient la base *Flora* du CBNBP. En effet, suite aux différents programmes d'atlas départementaux en Île-de-France, environ 1 500 000 données ont été collectées sur près de 35 000 polygones. Ces polygones sont géolocalisés et regroupent tous les taxons qui ont été observés sur un milieu théoriquement homogène. Ces polygones possèdent de nombreuses données de taxons inféodés à des milieux rares ou ponctuels qui peuvent être de bons indicateurs.

Seules les données postérieures à 1990, géolocalisées par un polygone cartographié sur SIG (Mapinfo), ont été sélectionnées pour cette étude.

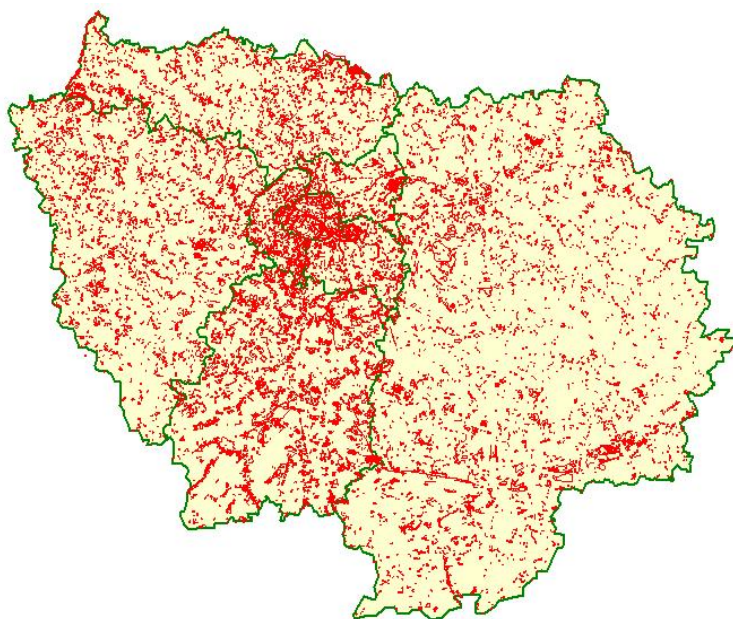


Figure 9 : Carte des données floristiques géolocalisées en Île-de-France. © CBNBP-MNHN, *Flora*.

Lors de la saisie des données floristiques dans la base de données *Flora* du CBNBP, plusieurs autres informations sont intégrées. Il s'agit de données stationnelles et de correspondances avec des codifications européennes. Dans le cadre de cette méthode, il a été choisi de retenir le code CORINE Biotopes car il reflète le grand type de milieu et offre un filtre intéressant.

2.2.2.2. Méthodologie

Une liste de syntaxon a été identifiée à partir du référentiel provisoire des végétations d'Île-de-France (Fernez, 2012). Ces syntaxons peuvent être subdivisés en deux groupes. Les premiers concernent les végétations patrimoniales peu ou pas recouvrantes ainsi que des végétations plus fréquentes mais toujours de faibles surfaces. Ces végétations sont largement sous-échantillonnées lors des cartographies sur de vastes territoires, et donc dans notre cas. Le second concerne des végétations plus communes et souvent très recouvrantes comme les forêts des *Quercus robur* - *Fagetea sylvaticae* Braun-Blanq. & J. Vlieger in J. Vlieger 1937 mais dont l'application dans la modélisation est présumée potentiellement intéressante (cf. partie. 1).

Une fois ces syntaxons sélectionnés, une liste floristique leur est alors associée. Cette liste floristique, réalisée à partir d'un travail bibliographique et des connaissances phytosociologiques régionales, a permis de sélectionner les espèces caractéristiques et différentielles régionales de chaque syntaxon cible. Les sources consultées sont les suivantes :

- Synopsis des groupements végétaux de la Bourgogne et de la Champagne-Ardenne (Royer et al, 2006) ;
- Guide des végétations des zones humides de la région Nord-Pas-de-Calais (Catteau *et al.*, 2009) ;
- Guide des végétations forestières et préforestières de la région Nord-Pas-de-Calais (Catteau *et al.*, 2010) ;
- Référentiel typologique du CBN de Brest (http://www.cbnbrest.fr/site/Refer_typo/habit0.php);
- Publications originelles où le syntaxon a été décrit lorsque cela été possible ;
- Données de la base de données *Habitat* du CBNBP;
- Clé des végétations d'Île-de-France (document interne).

Chaque syntaxon a donc une correspondance avec un ou plusieurs taxons indicateurs (cf. Tableau 1). Sur cette liste de taxons correspondants, si le nombre minimum de taxons présents dans un polygone est respecté, alors le polygone floristique est sélectionné et rattaché à ce syntaxon, ce fonctionnement est le même que pour les taxons caractéristiques d'une végétation donnée. Dans certains cas et pour affiner la recherche, deux listes de taxons indicateurs (Tableau 1, taxons A et taxons B) peuvent être associées et offrent ainsi un double filtrage.

La méthode de filtrage par des taxons indicateurs pour une association peut être couplée au filtrage de l'alliance correspondante. En effet, la requête au niveau de l'association utilise les taxons de la requête de l'alliance correspondante (Taxons A) en y ajoutant quelques taxons indicateurs (Taxons B) qui permettent de faire ressortir les différentes associations de l'alliance. Ainsi, lorsque les conditions n'étaient pas réunies pour faire ressortir un polygone à l'association, l'alliance pouvait toutefois être sélectionnée grâce au système de double liste de taxons. Cette organisation est calquée sur l'organisation hiérarchique du synsystème phytosociologique.

Ce système d'emboîtement n'est pas toujours été utilisé car, pour plusieurs alliances, il apparaît difficile de pouvoir être assez restrictif en utilisant qu'une seule liste de taxons indicateurs.

Dans de nombreux cas, un filtre avec le code CORINE Biotopes a également été intégré permettant d'affiner au maximum la sélection et la justesse des résultats.

Tableau 1 : Correspondances entre les syntaxons et les filtres (taxons indicateurs et code CORINE Biotopes)

Syntaxon	Nb A	Taxons A	Nb B	Taxons_B	Code CB
<i>Alyso alyssoidis - Sedion albi</i>	2 2 2 2	<i>Alyssum alyssoides</i> <i>Saxifraga tridactylites</i> <i>Veronica praecox</i> <i>Minuartia hybrida</i>			34
<i>Cerastietum pumili</i>	2 2 2 2	<i>Alyssum alyssoides</i> <i>Saxifraga tridactylites</i> <i>Veronica praecox</i> <i>Minuartia hybrida</i>	1	<i>Cerastium pumilum</i> <i>Hornungia petraea</i> <i>Bombycilaena erecta</i>	34
<i>Helianthemum apennini - Sedetum acris</i>	2 2 2 2	<i>Alyssum alyssoides</i> <i>Saxifraga tridactylites</i> <i>Veronica praecox</i> <i>Minuartia hybrida</i>	1	<i>Helianthemum apenninum</i>	34

Afin d'améliorer les résultats, d'autres filtres optionnels ont été intégrés :

- taxons négatifs : la présence de ces taxons élimine la sélection ;
- codes CORINE Biotopes négatifs : un relevé floristique renseigné avec un de ces codes élimine la sélection ;
- nombre maximal de taxons : au delà du nombre de taxons dans un même polygone, le polygone n'est pas retenu.

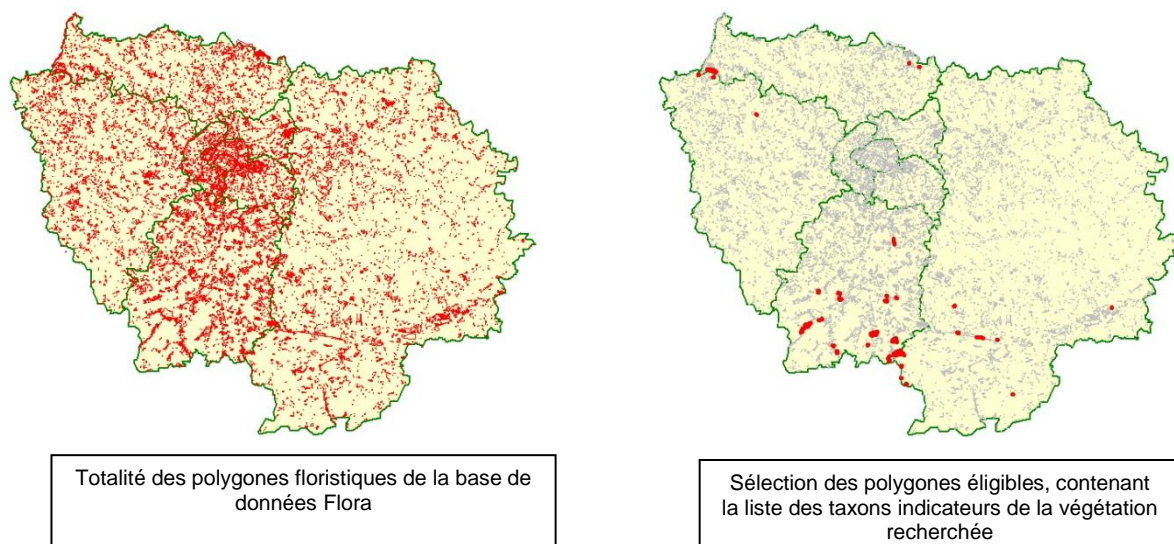
Les polygones de très grande superficie, trop hétérogènes (c'est-à-dire contenant de nombreux types de végétations différentes), ont aussi été écartés de la requête.

Lorsqu'il était considéré que la requête, pour une association donnée, n'aurait pas un résultat exploitable, elle était la plupart du temps élevée au niveau syntaxonomique supérieur (alliance), voir supprimée.

2.2.2.3. Résultats de la requête

Les résultats de l'extraction sont obtenus sous la forme d'une couche cartographique pour le logiciel Mapinfo et Editop. Ce dernier correspond au logiciel SIG utilisé lors du programme de cartographie des végétations d'Île-de-France.

Cette couche est complétée par un tableur Excel qui référence les polygones issus de la requête avec, pour chaque polygone, le syntaxon, la commune, l'identifiant *Flora* correspondant ainsi que les taxons indicateurs qui ont permis sa sélection.



Totalité des polygones floristiques de la base de données Flora

Sélection des polygones éligibles, contenant la liste des taxons indicateurs de la végétation recherchée

Figure 10 : Exemple de sélection de polygones de la base de données *Flora* filtré par les taxons indicateurs (*Alyssa alyssoides* - *Sedion albi* Oberd. & T. Müll. in T. Müll. 1961)

En 2012, 220 syntaxons ont été recherchés se répartissant en 122 associations, 21 sous-alliances, 75 alliances et 2 ordres phytosociologiques.

Ces 220 syntaxons ont fait ressortir un total de 6 000 polygones. Parmi ces polygones, environ 3 000 ne font ressortir qu'un seul syntaxon et peuvent donc être considérés comme les plus fiables.

2.2.2.4. Prospection sur le terrain

Au cours des prospections de terrain effectuées dans le cadre du programme de cartographie des végétations d'Île-de-France, plusieurs secteurs où des polygones issus de cette méthode sont ressortis, ont été prospectés. Aucun syntaxon ni secteur n'a été privilégié en raison du temps limité imparti au programme. De ce fait, il y a une forte hétérogénéité de pression d'observation *in-situ* entre les différents polygones qui ont été sélectionnés et donc entre les syntaxons correspondants.

2.3 – Bilan provisoire - Perspectives

2.3.1. Bilan

Les végétations qui ont bien fonctionné :

Les végétations à flore spécialisée semblent plutôt bien fonctionner. De nombreux polygones ressortent et le travail de terrain confirme la fiabilité de la requête. Cela concerne les végétations pionnières (*Isoeto durieui* – *Juncetea bufonii* Braun-Blanq. & Tüxen ex V. West., J. Dijk & Paschier 1946, *Littorelletea uniflorae* Braun-Blanq. & Tüxen ex V. West., J. Dijk & Paschier 1946, *Sedo albi* – *Scleranthetea perennis* Braun-Blanq. 1955...) et les végétations aquatiques (*Lemnetea minoris* Tüxen ex O. Bolòs & Masclans 1955, *Potametea pectinati* Klika in Klika & V. Novák 1941).

En effet, ces groupements sont caractérisés par un faible nombre d'espèces, mais souvent très inféodées à un type de milieu particulier. C'est pourquoi ces polygones ressortent facilement et sont plutôt fiables. De plus, les relevés d'inventaire floristique concernant ces milieux particuliers sont souvent réalisés sur un milieu homogène puisque ces végétations sont assez différentes de celles qui les entourent, permettant de favoriser la réussite de la méthode.

La méthode fonctionne donc bien pour ces végétations ponctuelles et souvent d'intérêt patrimonial (habitats d'intérêt communautaire en particulier) et permet de les détecter lors de cartographies surfaciques.

Les végétations aux résultats mitigés :

Les résultats pour les pelouses des *Festuco valesiacae* – *Brometea erecti* Braun-Blanq. & Tüxen ex Braun-Blanq. 1949 sont assez intéressants. Ces végétations d'un grand intérêt patrimonial et souvent de faible surface ont donné des résultats plutôt hétérogènes. Les pelouses sablo-calcaires ont été globalement bien détectées, car les taxons indicateurs choisis sont assez exclusifs de ces milieux. Les polygones qui sont ressortis sont donc assez fiables mais ils sont peu nombreux. En revanche, les pelouses calcicoles (*Mesobromion erecti* (Braun-Blanq. & Moor 1938) Oberd. 1957 et *Xerobromion erecti* (Braun-Blanq. & Moor 1938) Moravec in Holub, Hejný, Moravec & Neuhäusl 1967) n'ont pas donné de résultats vraiment concluants. La plupart correspondent, soit à des pelouses sablo-calcaires, soit à des ourlets calcicoles avec lesquels les pelouses calcicoles partagent de nombreux taxons en commun. Seules les pelouses avec des taxons très localisés dans la région et exclusifs de ces végétations (*Linum leonii* ou *Astragalus monspessulanus* par exemple) ont bien fonctionné.

Les résultats pour les pelouses sont à temporiser car, dans le cadre de ce programme, la plupart des pelouses sont échantillonnées sur le terrain ce qui limite l'intérêt de la méthode pour ces végétations.

Les végétations qui n'ont pas fonctionné :

Pour certaines végétations, la méthode n'a pas fonctionné. C'est le cas des végétations à taxons plus ou moins ubiquistes, comme les prairies au sens large des *Arrhenatheretea elatioris* Braun-Blanq. ex

Braun-Blanq. *et al.* 1952, des *Agrostietea stoloniferae* Oberd. 1983 et des *Molinio caeruleae* – *Juncetea acutiflori* Braun-Blanq. 1950, où les résultats à l’alliance ou à l’association ne sont pas fiables. Seules les végétations dont les taxons indicateurs sont très rares et exclusifs de ces végétations dans la région, sont plutôt bien détectées mais présentent très peu de polygones. C’est notamment le cas des prairies de l’*Oenanthion fistulosae* B. Foucault 2008 (4 polygones sur toute la région).

Certaines végétations linéaires ou ponctuelles qui sont régulièrement en lisières d’autres végétations de surfaces plus importantes, ne sont que très rarement individualisées dans les inventaires floristiques. C’est le cas des ourlets (*Galio aparines* – *Urticetea dioicae* H. Passarge ex Kopecký 1969, *Melampyro pratensis* – *Holcetea mollis* H. Passarge 1994 et *Trifolio medii* – *Geranietea sanguinei* T. Müll. 1962) et des manteaux forestiers (*Crataego monogynae* – *Prunetea spinosae* Tüxen 1962 et *Cytisetetea scopario-striati* Rivas Mart. 1975). Le résultat pour ces végétations est donc très peu fiable et correspond la plupart du temps à des polygones de forêts. De plus, ces groupements ont de nombreux taxons en commun, limitant encore la fiabilité des résultats entre eux.

Enfin, les milieux forestiers, qui représentent le milieu naturel le plus important de la région, n’ont pas donné de résultats intéressants. Le test avait été réalisé au niveau de l’association pour l’alliance du *Carpino betuli* – *Fagion sylvaticae* Bœuf, Renaux & J.-M. Royer *in* Bœuf 2011 afin d’essayer de localiser les différents groupements de ces milieux. Le choix a été de faire ressortir le maximum de polygones en incluant des taxons indicateurs trop larges pour chaque association. La plupart de ces taxons peuvent se retrouver dans les différentes associations forestières, leur abondance relative permettant de les différencier.

Le manque de fiabilité pour ce type de milieux a été accentué par le fait que les relevés floristiques forestiers sont régulièrement à cheval sur plusieurs communautés floristiques différentes.

Les végétations non testées :

Les végétations rudérales (*Agropyreteae pungentis* Géhu 1968, *Artemisietea stoloniferae* W. Lohmeyer, Preising & Tüxen *ex* von Rochow 1951, *Sisymbrieteteae officinalis* Korneck 1974) n’ont pas été testées car elles ne présentent pas d’intérêt particulier et que leur composition rend les résultats de la méthode très peu fiables. En effet, ces végétations sont souvent composées d’un grand nombre de taxons issus de milieux variés, les taxons caractéristiques des friches se retrouvant régulièrement dans d’autres végétations.

2.3.2. Limites de la méthode

Le bilan a mis en évidence que, dans la plupart des cas, une recherche au niveau de l’association semble peu fiable. En effet, et ce malgré l’application de tous les filtres cités, un certain nombre de polygones floristiques s’allument simultanément pour plusieurs associations qui n’appartiennent pas à la même classe phytosociologique dans certains cas.

Par ailleurs, une recherche trop restrictive augmente la fiabilité mais fait également baisser le nombre de polygones sélectionnés et quelquefois seul un ou deux polygones sont retenus, limitant l’intérêt de

la méthode. Au contraire, une recherche plus large va faire ressortir un plus grand nombre de polygones mais la plupart seront erronés. La difficulté est donc de trouver le bon rapport entre niveau de recherche et de filtre, et spécificité et nombre de polygones sélectionnés.

De ce fait, il semble que le niveau de recherche ne puisse se faire en dessous de l'alliance ou de la sous-alliance.

Le système emboîté alliance-association ne fonctionne pas bien non plus et les taxons indicateurs de l'association ne peuvent pas reprendre tous les taxons indicateurs de l'alliance en y ajoutant quelques taxons caractérisant le niveau association.

Enfin et c'est le principal problème, l'hétérogénéité des relevés d'inventaire floristique, parfois effectués sur des milieux aux conditions homogènes, mais la plupart du temps, afin de gagner du temps dans la réalisation des relevés, sur des surfaces importantes couvrant de nombreux types de végétations au sein d'un même polygone, cette méthode permettant d'inventorier le plus grand nombre d'espèces possible dans le temps imparti à la réalisation de l'inventaire. Sur des tels polygones, notre méthode de détection des végétations se heurte ainsi au problème de la multiplicité et de la complexité des végétations intrinsèques, rendant le résultat souvent non directement exploitable malgré l'application des filtres complémentaires, parfois même erroné.

2.3.3. Perspectives

Cette méthode sera, de nouveau, mise en place dans le cadre du programme de cartographie des végétations d'Île-de-France lors de la campagne 2013.

Les enseignements de l'année 2012 permettront d'améliorer la méthode pour atteindre la meilleure fiabilité possible. Pour cela, seules les végétations patrimoniales et/ou ponctuelles largement sous-prospectées seront étudiées. Le choix des taxons indicateurs sera également affiné en tenant compte de ce qui a plutôt bien fonctionné jusqu'à maintenant.

Pour améliorer les résultats, des informations environnementales pourraient être couplées avec la méthode de localisation de végétations à partir de données floristiques géoréférencées. Cela permettrait d'ajouter de nouveaux filtres à la méthode et augmenterait ainsi fortement sa fiabilité. Ces filtres environnementaux pourraient concerner les variables classiques, déterminantes dans la distribution des végétations, telles que la pente, l'exposition, la géologie ou encore la petite région naturelle à un niveau plus intégrateur.

Conclusion

Les deux méthodes présentées dans ce document ont été initiées dans des contextes et des objectifs différents. La modélisation en Seine-et-Marne a pour objectif de déterminer des végétations non inventoriées sur le terrain, contrairement à la localisation des végétations qui, à partir de données floristiques géoréférencées, a pour objectif de mettre en évidence des végétations patrimoniales inventoriées en terme de listes d'espèces mais sous-prospectées en terme de végétations. Les données utilisées dans le cadre de ces deux méthodes ne sont donc pas du tout similaires. Dans la modélisation, les données Végétation sont directement exploitées pour attribuer des syntaxons à des polygones issus d'une cartographie des milieux photo-interprétée qu'est l'Ecomos, alors que dans la localisation des végétations, ce sont des données de terrain issues de relevés floristiques qui ciblent précisément certains types de végétations.

Pourtant, les deux méthodes sont complémentaires puisqu'elles participent à la finalisation de la cartographie des végétations d'Île-de-France. Elles ne concernent pas les mêmes types de végétations, sauf dans certains cas où les résultats ont montré qu'une méthode peut afficher des meilleurs taux de réussite que l'autre (ex : pelouses calcicoles). Il est alors possible de choisir la méthode en fonction des milieux et des tailles de polygones à renseigner. Quoiqu'il en soit, elles seront à coupler afin de modéliser le plus fidèlement possible les végétations de la région au niveau de l'alliance.

Les perspectives dans le cadre de Carhab sont nombreuses, autant pour la méthode de modélisation surfacique que pour celle de détection des végétations par les données d'inventaire floristique :

- Pour la modélisation surfacique sur le fond blanc de Carhab, avec l'apport de nouvelles variables environnementales spatialisées et l'utilisation du même jeu de variables que celui du fond écologique, travail en cours d'élaboration par le laboratoire EVS-ISTHME de l'université Jean Monnet de Saint-Etienne (Etlicher *et al.*, 2013),
- Pour la localisation des végétations patrimoniales ponctuelles que la cartographie phytosociologique et dynamico-caténale au 1/25 000ème, telle que prévue dans Carhab, ne pourra que trop rarement détecter. Les données récentes et géolocalisées d'inventaire floristique existent maintenant sur tout le territoire français, en particulier dans les bases de données des conservatoires botaniques nationaux. La méthode de détection proposée et testée ici pourrait ainsi être déployée à l'échelle nationale. Elle nécessiterait l'identification et éventuellement la hiérarchisation des végétations ciblées selon des critères définis, l'élaboration des listes d'espèces indicatrices pour chacune d'entre elles, le paramétrage du filtre espèces et des filtres complémentaires éventuels. La détection de végétations fortement patrimoniales pourrait ainsi compléter directement la cartographie surfacique par des données ponctuelles à fort enjeu, mais également orienter les prospections terrain sur ces secteurs lors du déploiement de la cartographie surfacique au 1/25 000ème.

Bibliographie

- Azuelos L. & Renaud H., 2009. Notice explicative des cartographies de végétations produites par le CBNBP pour le CG77. Qualité et exploitation des données récoltées, Septembre 2009. Conservatoire Botanique National du Bassin Parisien, 33 p.
- Bossard M., Feranac J. & Otahel J., 2000. CORINE land cover technical guide. Addendum 2000. European Environment Agency.
- Catteau E., Duhamel F., Baliga M.F., Basso F., Bedouey F., Cornier T., Mullié B., Mora F., Toussaint B. & Valentin B., 2009. Guide des végétations des zones humides de la région Nord-Pas de Calais. Centre régional de Phytosociologie agréé Conservatoire botanique national de Bailleul, Bailleul, 632 p.
- Catteau E., Duhamel F., Cornier T., Farvaques C., Mora F., Delplanque S., Henry E., Nicolazo C., Valet J.-M., 2010. Guide des végétations forestières et préforestières de la région Nord-Pas de Calais. Centre régional de Phytosociologie agréé Conservatoire botanique national de Bailleul, Bailleul, 526 p.
- Conseil Régional de Seine-et-Marne (CG77), 2007. Atlas des paysages de Seine-et-Marne – CG77 & CAUE 77, 273 p.
- Direction Régionale de l'Agriculture et de la Forêt (D.R.A.F.), 1992. Île-de-France. Tableaux et graphiques. Agreste - La statistique agricole, Paris, 62 p.
- Dupias G. & Rey P., 1985. Document pour un zonage des régions phyto-écologiques. CNRS, Toulouse, 39 p. + 1 carte.
- Etilcher B., Mazagol P.-O., Sacca C. & Fayeaux F., 2013. Programme national CARHAB. Recherches concernant la cartographie nationale des habitats naturels et semi-naturels. La construction du fond écologique. Rapport d'étape. EVS-ISTHME, Univ. Jean Monnet Saint-Etienne.
- Fernez T., version provisoire 2012. Synopsis phytosociologique des groupements végétaux d'Île-de-France. Conservatoire botanique national du Bassin parisien.
- Filoché S., Perriat F., Hendoux F. & Moret J., 2010. Atlas de la flore sauvage de Seine-et-Marne. Conservatoire Botanique National du Bassin Parisien, Librairie des Musées /Conseil général de Seine-et-Marne, 688 p.
- Heymann Y., Steenmans C., Croisille G. & Bossard S., 1994. CORINE land cover: Technical guide. Luxembourg: Office for Official Publications of the European Communities.

Institut d'Aménagement et d'Urbanisme (IAU), 2005. ECOMOS 2000 ou la cartographie détaillée des milieux naturels en Île-de-France. Note rapide sur l'environnement, n° 388, Juin 2005, 6 p. <http://www.iau-idf.fr/detail-dune-etude/etude/ecomos-2000-ou-la-cartographie-detaillee-des-milieux-na.html>

Institut d'Aménagement et d'Urbanisme (IAU), 2005. Les derniers résultats du MOS 2003. Note rapide Mode d'occupation du sol, n° 387, Juin 2005, 6 p. <http://www.iau-idf.fr/detail-dune-etude/etude/les-derniers-resultats-du-mos-2003.html>

Lafon P., 2012. Essai de localisation géographique de végétations à partir de l'exploitation de données floristiques géolocalisées : méthodologie, bilan et perspectives. Conservatoire Botanique National du Bassin Parisien, 18 p.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Royer J.-M., Felzines J.-C., Misset C. & Thévenin S., 2006. Synopsis commenté des groupements végétaux de la Bourgogne et de la Champagne-Ardenne. Bull. Soc. Bot. Centre-Ouest, Nouvelle série, N.S. 25, 394 p.

Annexe 1 : Types d'Ecomos (CLC6)

CLC6	Nom_CLC6
231150	prairie complantée
231121	prairie mésophile avec arbrisseaux
231131	prairie mésophile avec arbres
231113	prairie mésophile "propre" sous ligne HT ou THT
231112	prairie mésophile "propre" en forêt
231122	prairie mésophile avec arbrisseaux, en forêt
231132	prairie mésophile avec arbres, en forêt
231140	prairie à gibier
231114	prairie mésophile "propre", le long d'infrastructures
231111	prairie mésophile "propre"
231221	prairie humide avec arbrisseaux
231231	prairie humide avec arbres
231214	prairie humide "propre", en bassin de décantation
231212	prairie humide "propre", en forêt
231222	prairie humide avec arbrisseaux, en forêt
231232	prairie humide avec arbres, en forêt
231213	prairie humide "propre", liée à des infrastructures de transport
231211	prairie humide "propre"
312150	coupe en résineux
311750	coupe en peupleraie
311152	coupe en forêt de feuillus xéro à mésophile, sous ligne HT ou THT
311151	coupe en forêt de feuillus xéro à mésophile
324113	formations herbacées et boisées sous ligne HT ou THT
324111	formations herbacées et boisées liées aux infrastructures de transport
324112	formations herbacées et boisées liées aux terrains de sport
324114	formations herbacées et boisées liées aux autres installations (ZI, ...)
332201	sol à nu d'ancienne carrière
332202	sol à nu en dépendance d'infrastructures
332100	affleurement rocheux ou sableux
222200	verger non entretenu
231340	friche sur ancien verger
231320	friche herbacée avec arbrisseaux
231330	friche herbacée avec arbres
231400	friche humide sur ancienne cressonnière
231313	friche purement herbacée, sous ligne HT ou THT
231312	friche purement herbacée, le long d'infrastructures de transport
231311	friche purement herbacée
311112	feuillus denses xéro à mésophile, bas, sous ligne HT ou THT
311113	feuillus denses xéro à mésophile, aux abords d'infrastructures
311111	feuillus denses xéro à mésophile
311210	jeunes feuillus à callune
311123	boisement de jeunes feuillus xéro à mésophile, sous ligne HT ou THT
311122	boisement de jeunes feuillus xéro à mésophile, aux abords d'infrastructures
311310	jeunes feuillus sur platière gréseuse
311121	boisement de jeunes feuillus xéro à mésophile
311220	feuillus clairs à callune
311322	feuillus clairs sur chaos gréseux

CLC6**Nom_CLC6**

311135 feuillus clairs xéro à mésophile, sous ligne HT ou THT
311321 feuillus clairs sur platière gréseuse
311132 régénération de feuillus xéro à mésophile
311134 régénération en feuillus xéro à mésophile, décimés par la tempête
311400 feuillus clairs sur sable
311133 feuillus clairs xéro à mésophile, décimés par la tempête
311131 feuillus clairs xéro à mésophile
311143 forêt de feuillus xéro à mésophile, avec moins de 10% d'arbres, sous ligne HT ou THT
311142 forêt de feuillus xéro à mésophile, décimée par la tempête et encore éclaircie depuis
311141 forêt de feuillus xéro à mésophile, ravagée par la tempête, avec moins de 10% d'arbres
311611 forêt marécageuse dense
311612 forêt marécageuse dense sous ligne HT ou THT
311510 forêt humide dense
311522 forêt humide dense, sous ligne HT ou THT
311620 forêt marécageuse jeune
311521 forêt humide jeune
311630 forêt marécageuse claire
311532 forêt humide claire, suite à la tempête
311531 forêt humide claire
311540 forêt humide rasée
312110 résineux dense
312210 jeunes résineux sur callune
312122 jeunes résineux, sous ligne HT ou THT
312121 jeunes résineux
312220 résineux clairs sur callune
312311 résineux clairs sur chaos
312400 résineux clairs sur pelouse
312312 résineux clairs sur platière gréseuse
312132 résineux clairs, suite à la tempête
312131 résineux clairs
312140 résineux rasés
313210 forêt mélangée dense à callune
313110 forêt mélangée mésophile dense
313120 forêt mélangée mésophile jeune
313220 forêt mélangée claire à callune
313312 forêt mélangée mésophile claire sur chaos gréseux
313311 forêt mélangée mésophile claire sur platière gréseuse
313132 forêt mélangée mésophile claire, suite à la tempête
313131 forêt mélangée mésophile claire
313140 forêt mélangée mésophile claire
311710 peupleraie dense
311722 peupleraie jeune sur mégaphorbiaie
311721 peupleraie jeune
311732 peupleraie claire, suite à la tempête
311731 peupleraie claire
311740 peupleraie rasée
242210 jardin familial en friche
242220 jardin de l'habitat en friche
324212 friche arbustive basse sur ancienne terre agricole, sous ligne HT ou THT
324213 friche arbustive basse sur ancienne terre agricole, le long d'infrastructure

324211 friche arbustive basse sur ancienne terre agricole, en terrain ouvert
324300 verger en friche
324220 friche avec arbres sur ancienne terre agricole
243100 culture à gibier + parcelle boisée
321120 pelouse calcaire avec arbrisseaux
321220 pelouse sur sable avec arbrisseaux
321130 pelouse calcaire avec arbres
321400 pelouse en carrière
321320 pelouse sur chaos gréseux
321112 pelouse calcaire herbacée, sous ligne HT ou THT
321113 pelouse calcaire herbacée, liée à des infrastructures de tempête
321310 pelouse sur platière gréseuse
321210 pelouse herbacée sur sable
321111 pelouse calcaire herbacée
322420 autres landes ou landes indifférenciées, avec arbres
322100 lande à callune
322220 lande sur chaos gréseux
322412 autres landes ou landes indifférenciées, sous ligne HT ou THT
322300 lande humide
322210 lande sur platière gréseuse
322411 autres landes ou landes indifférenciées
324700 fourrés avec callune
324620 fourrés sur chaos
324232 friche très dense avec arbres sur ancienne terre agricole, sous ligne HT ou THT
324401 prébois calcicole
324402 prébois calcicole, sous ligne HT ou THT
324600 fourrés sur platière
324500 prébois sur sable
324231 friche très dense avec arbres sur ancienne terre agricole, en terrain ouvert
333220 végétation clairsemée sur chaos rocheux
333102 végétation clairsemée, sous ligne HT ou THT
333210 végétation clairsemée sur platière gréseuse
333101 végétation clairsemée
411620 végétation humide, avec arbustes
411630 végétation humide, avec arbres
411300 magnocaricaie
411613 végétation humide en bassin de décantation
411100 grève d'étang
411612 végétation humide en bassin d'infrastructure
411612 végétation humide en bassin de décantation
411400 mégaphorbiaie
411200 roselière
411500 zone marécageuse avec saules
411611 végétation humide basse
412000 tourbière
512120 plan d'eau permanent libre, avec quelques arbres
512112 bassin de décantation, en eau libre
512400 mouillère
512300 plan d'eau avec nénuphar
512200 plan d'eau avec végétation aquatique
512111 plan d'eau permanent libre

Annexe 2 : Script R

```
##### Script Extrapolation #####
```

```
### Appel des données ###
```

```
library(RODBC)
```

```
library(sqldf)
```

```
conn <- "D:/Private/Île-de-France/Seine-et-Marne/SM_2012/Habitats20120314.mdb"
```

```
channel <- odbcConnectAccess(conn)
```

```
tables <- sqlTables(channel)
```

```
EcoDon_2012 <- sqlFetch(channel,"Table_Ecomos_Donnees_2012_rogne") ; dim(EcoDon_2012) # 19251 10
```

```
Hab_2012 <- sqlFetch(channel,"Table_Habitat_Donnees_2012_max") ; dim(Hab_2012) # 15477 22
```

```
odbcClose(channel)
```

```
Hab2_2012 <- unique(Hab_2012[,c(1:16,18,22)]) ; dim(Hab2_2012) # 6651 18
```

```
Habvue <- Hab2_2012
```

```
### Modèle par petite région et par CLC6 ###
```

```
reg = unique(as.vector(Habvue$regnat))
```

```
for (ir in 1:length(reg)) {
```

```
  print(reg[ir])
```

```
  HabvueS = Habvue$Clc6[which(Habvue$regnat==reg[ir])] # levels(Habvue$regnat)
```

```
  print(length(HabvueS))
```

```
  ecom = unique(HabvueS)
```

```
  fout = "Out120612.txt" # fichier de sortie
```

```
  lvar = c(11,12,13,14,15,16) # numéro des variables à utiliser dans l'analyse
```

```
  for (ieco in 1:length(ecom)) {
```

```
    HabvueS = Habvue[which(Habvue$regnat==reg[ir] & Habvue$Clc6==ecom[ieco]),]
```

```
    HabvueS$Nv_Syntax_gene = as.factor(as.vector(HabvueS$Nv_Syntax_gene))
```

```
    cat(ieco,"",length(ecom)," - ",nrow(HabvueS),"id_habitats, ",nlevels(HabvueS$Nv_Syntax_gene),"syntaxons  
pour",ecom[ieco],"\n")
```

```
    j <- which(is.na(HabvueS[,lvar]),arr.ind=TRUE)
```

```
    if (nrow(j) > 0) HabvueS <- HabvueS[-j[,1],]
```

```
    sup = EcoDon_2012[which(EcoDon_2012$regnat==reg[ir] & EcoDon_2012$Clc6==ecom[ieco]),c(1,5:10)]
```

```
    cat(nrow(sup),"sites à affecter pour",ecom[ieco],"\n")
```

```
    if (nrow(sup) > 0 & nrow(HabvueS[,lvar]) > 0) {
```

```
      if (nrow(HabvueS) == 1 | sum(cov(HabvueS[,lvar]))==0) {
```

```

j = which.max(HabvueS$Surface)

cat(reg[ir],ecom[ieco],as.vector(HabvueS$id_habitat[j]),HabvueS$Surface[j],sup[1,1],as.vector(HabvueS$Nv_Synt
ax_gene[j]),"*\n",file=fout,sep=";",append=TRUE)
} else {
acp = dudi.pca(HabvueS[,lvar],scannf=FALSE,nf=length(lvar))
ax = c(1,ifelse(length(acp$eig)==1,1,2))
plot(acp$li[,ax],pch=19,cex=.8,col=as.numeric(HabvueS$Nv_Syntax_gene),main=reg[ir],sub=ecom[ieco])
acpsup = suprow(acp,sup[,2:7])
for (k in 1:nrow(acpsup$li)) {
j = as.matrix(dist(rbind(acpsup$li[k,],acp$li)))[-1,1]
j = which(j%in%min(j)) #
n = which.max(HabvueS$Surface[j])
j = j[n]
points(acpsup$li[k,ax],col=as.numeric(HabvueS$Nv_Syntax_gene[j]))

cat(reg[ir],ecom[ieco],as.vector(HabvueS$id_habitat[j]),HabvueS$Surface[j],sup[k,1],as.vector(HabvueS$Nv_Synt
ax_gene[j]),"\n",file=fout,sep=";",append=TRUE)
cat (k ,"/",nrow(acpsup$li),j,"\n" )
} }
}

```