



Protocole d'échantillonnage physionomique pour la télédétection des végétations : version 1

Samuel Alleaume, C. Corbane, M. Isenmann, Vincent Thiérion, C. Jacqueminet

► To cite this version:

Samuel Alleaume, C. Corbane, M. Isenmann, Vincent Thiérion, C. Jacqueminet. Protocole d'échantillonnage physionomique pour la télédétection des végétations : version 1. irstea. 2013, pp.12. hal-02599606


HAL Id: hal-02599606

<https://hal.inrae.fr/hal-02599606>

Submitted on 16 May 2020

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.



Protocole d'échantillonnage physionomique pour la télédétection des végétations. –version 1-

2013

Samuel ALLEAUME

Christina CORBANE

Marc ISENMANN

Vincent THIERION

Christine JACQUEMINET

Irstea- UMR TETIS

Rapport réalisé dans le cadre du projet
CARHAB

Pour mieux
affirmer
ses missions,
le Cemagref
devient Irstea



Table des matières

1.	INTRODUCTION	3
2.	DEFINITIONS	3
3.	CHOIX DE LA NOMENCLATURE.....	5
4.	SEGMENTATION ADOPTEE	7
5.	PRODUIT ATTENDU	7
6.	VALORISATION DES CARTES EXISTANTES :	8
7.	CREATION DE NOUVEAUX ECHANTILLONS.....	11
8.	REFERENCES.....	12

1. INTRODUCTION

La télédétection est une approche incontournable lorsqu'il s'agit de cartographier les végétations naturelles sur de vastes territoires. Les techniques de classification automatisées d'images nécessitent cependant d'avoir recours à des échantillons représentatifs des classes à cartographier. Ce besoin s'exprime à deux étapes de la classification d'image : en amont durant la phase de calibration et en aval pour la validation.

Cependant, le passage d'une observation de terrain vers une donnée exploitable en termes de classification d'image n'est pas toujours évident. L'expérience au cours du projet Carhab a montré les difficultés rencontrées face aux remontées de terrain :

- difficultés spatiales : données ponctuelles non compatibles avec les surfaces à cartographier, incohérences d'échelles entre les documents fournis et les exigences Carhab (1 :25 000).
- difficultés temporelles : les cartographies fournies ne sont pas nécessairement récentes ou en cohérence avec les données satellites
- difficultés de nomenclatures : les cartes ont été faites dans une typologie qui est difficile à retraduire dans la nomenclature physionomique

Par conséquent, ce présent document propose une méthode afin d'harmoniser les acquisitions de terrain afin qu'elles soient compatibles avec les méthodes de télédétection.

Ce protocole nécessite de bonnes compétences en géomatique (SIG), une bonne connaissance des données mobilisées et une expertise sur les nomenclatures utilisées pour les cartographies de la végétation. Il peut donc impliquer la contribution de deux acteurs différents dont les compétences sont complémentaires. A cette étape, la spécialisation en télédétection n'est pas requise. Cependant, puisque ce travail va être exploité pour un traitement par télédétection, quelques notions dans ce domaine sont redéfinies.

2. DEFINITIONS

Cette série de définitions a pour objectif de montrer l'importance de créer un échantillonnage de qualité qui servira comme sites d'entraînement afin de classer la végétation ouverte de basse et de haute altitude par télédétection. Ces échantillons contribuent également à la validation de la classification. Cet échantillonnage doit être réalisé sur la base d'une segmentation adaptée à l'échelle de travail ainsi qu'à la nomenclature adoptée.

Sites d'entraînement

Traditionnellement, l'étalonnage de toute classification par télédétection doit s'appuyer sur des échantillons de terrain qui correspondent à des zones spectralement homogènes de la classe d'intérêt. Ces zones homogènes sont appelées des sites d'entraînement. Ces derniers peuvent être sélectionnés d'après des cartes existantes, des données d'observation au sol, des photos aériennes stéréos interprétées ou encore par combinaison de plusieurs approches. Les zones d'entraînement doivent posséder les caractéristiques suivantes : être représentatives du type de végétation relevé ; suffisamment vastes pour offrir une bonne représentation statistique de la classe ; bien distribuées spatialement pour minimiser l'effet des variations locales sur la classification.

Segmentation

La segmentation permet de délimiter des objets (ou segments) correspondant à des zones homogènes de l'image en fonction de critères de couleur (valeur spectrale) et de forme. C'est la base de la classification orientée-objets. La taille des objets est également dépendante du choix du facteur d'échelle adopté dans le processus de segmentation. Plus ce paramètre est élevé plus les objets sont grands. Il est possible de réaliser des segmentations multirésolution, c'est-à-dire à plusieurs facteurs d'échelle. Ces différents niveaux de segmentation sont emboîtés, ce qui permet de réaliser des classifications basées sur différents niveaux d'échelle. La segmentation crée un lien conceptuel avec une perception humaine en essayant d'imiter la capacité de perception du photo-interprète (Lang et al., 2009; Weinke et al., 2008).

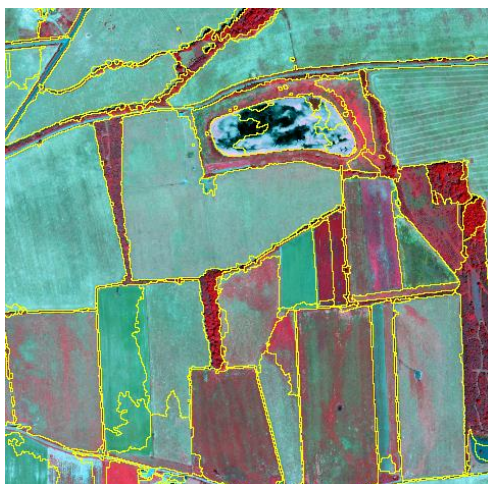


Figure 1. Exemple de segmentation IGN (en jaune) sur le fond IRC IGN.

Classification

La classification consiste à attribuer une classe à des objets ou des pixels en fonction des variables qui les caractérisent. Les types de classification qui nécessitent des sites d'entraînement pour un apprentissage sont de deux familles : les classifications classiques supervisées et la méthode experte. Ces deux approches présentées ici distinctement peuvent être utilisées de façon complémentaire dans le cadre d'un projet.

Classification supervisée : l'analyste s'entraîne sur ces sites afin de retrouver des objets comparables sur l'ensemble de l'image étudiée. L'algorithme de classification ou classificateur doit être calibré à l'aide de sites d'entraînement correspondant à des zones homogènes et parfaitement représentatives de chacune des classes à cartographier. Cette méthode est illustrée à travers la classification par plus proche voisin, fréquemment utilisée en approche orientée-objet. La réponse moyenne de chaque objet de l'image est calculée et comparée statistiquement à celle des échantillons et l'objet est rangé dans la classe ayant l'échantillon le plus proche de l'objet selon les attributs retenus (Figure 2).

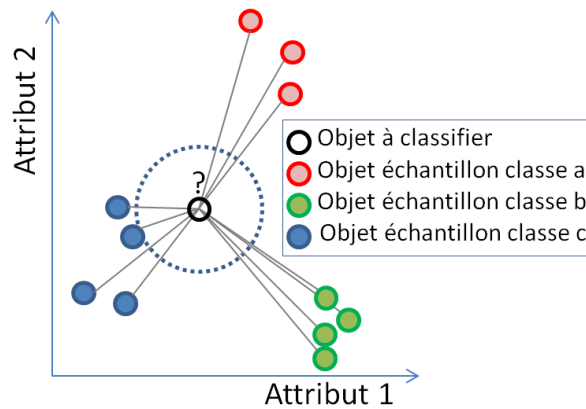


Figure 2. Principe de la classification par plus proche voisin.
Ici la classe C sera attribuée à l'objet à classer puisque sa distance en est plus proche dans l'espace des attributs 1 et 2.

L'avantage de cette méthode réside dans le fait qu'il n'y a pas de choix *a priori* des paramètres (ou attributs) qui vont permettre de séparer les classes à détecter, les différentes bandes spectrales de l'image pouvant être utilisées.

Classification par méthode experte : méthode encore connue sous le nom de règles de décision ou encore règles de connaissance ou à base de règles de classification. Il s'agit de combiner un certain nombre de règles pour associer chaque objet à une typologie définie. Une règle consiste à établir des seuils de valeurs des attributs (couleur, géométrie, contexte) choisis en fonction de l'observation et des connaissances du télédéacteur. La valeur des seuils se fait soit de manière empirique ou soit assistée en utilisant des techniques statistiques. En utilisant cette approche, l'analyste peut progressivement et hiérarchiquement classer l'image et ajuster itérativement les règles de décision pour les classes les plus problématiques. Cette méthode garantit a priori une meilleure reproductibilité de la méthode.

Validation

La validation consiste à établir si la représentation des classes données par la télédétection est en accord avec les données collectées sur le terrain. Pour cela, on a recours à l'utilisation des matrices de confusion. Elles permettent de quantifier la précision avec laquelle les différents thèmes ont été classifiés et, d'autre part, de définir quels sont les thèmes présentant des confusions thématiques (Girard et Girard, 2010). Elle est obtenue en comparant les données classées (en lignes) avec des données de référence (en colonnes) qui doivent être différentes de celles ayant servi à réaliser la classification.

3. CHOIX DE LA NOMENCLATURE

La nomenclature utilisée doit rendre compte à la fois des besoins de l'utilisateur mais aussi des capacités et des limites de l'outil de télédétection. Par conséquent la construction de la nomenclature doit s'établir au travers d'un dialogue constant entre les deux acteurs.

La nomenclature proposée pour la classification des milieux ouverts est en accord avec celle adoptée lors d'une réunion « Transect Isère » du 26 mars 2013 réunissant tous les experts en télédétection du projet Carhab (Tableau 1). Il s'agit d'une nomenclature purement physionomique dans la mesure où la végétation est décrite en termes de structure et de

production végétale. Cette nomenclature est suffisamment généraliste pour qu'elle puisse s'adapter à tous types de formations végétales, quel que soit la zone phytogéographique étudiée. Il est donc important de décrire la ou les végétation(s) inclue(s) dans chaque type physionomique sous la forme d'une notice fournie avec la carte.

Tableau 1. Nomenclature physionomique pour les milieux ouverts. Adoptée suite à la réunion « Transect Isère » du 26 mars 2013.

Niveau 1	Niveau 2	Niveau 3
Strates	Objectif prioritaire	basé sur des critères écologiques et temporels
Eau	surface en eau (BD Topo IGN)	
Minéral	Eboulis – sol nu Herbacé – minéral	Paroi (végétalisée ou non) Sol nu Eboulis stérile ou végétalisé Chaos de blocs
Herbacée	Herbacée à phytomasse faible Herbacée à phytomasse moyenne Herbacée à phytomasse forte	Prairie permanente Fauchée, pâturée Hydromorphie (bas marais, etc.)
Ligneux bas < 5 m	Ligneux bas rocailleux Ligneux bas mixte Ligneux bas dense Fourré – Fruticée	
Ligneux haut	Forêt lâche	

Niveau 1 : Strate naturelle ou semi-naturelle

Herbacée : zone enherbée plus ou moins permanente. Elle correspond à une prairie ou une pelouse.

Ligneux bas : correspond à la strate sub-arborée (<5m). Souvent associés aux milieux de landes ou à des milieux en voie de recolonisation.

Ligneux haut : zones arborées comprenant essentiellement les forêts et les haies. Cette information est prise comme un masque pour cette étude, puisque qu'elle ne concerne pas les milieux ouverts. Le milieu forestier est masqué par la classe « forêt fermée » de la BD Forêt. Les haies sont extraites de la Bd Végétation (BD TOPO).

Autre : zones artificialisées (zones urbaines, routes), cultures...

Niveau 2 :

Phytomasse. S'applique à la strate herbacée. Il s'agit d'une indication sur le maximum annuel de production chlorophyllienne. Ce caractère est décliné en 3 niveaux de productivité en phytomasse : faible, moyenne et forte. Cet indicateur est intégrateur de la hauteur et de la densité de la strate herbacée.

Densité. S'applique aux ligneux. Pour un milieu ouvert on utilisera le qualificatif de mixte puisque la strate herbacée est encore bien représentée. A l'opposé, les ligneux à forte

densité sont nommés denses. Enfin, le type « Ligneux bas rocaillieux » concerne des mosaïques de landes, pelouses et minéral. La strate **fourré – fruticée** est une formation moyennement haute et également dense. Une grande part de confusion est possible avec les ligneux bas denses, puisque le caractère de hauteur de végétation n'est pas déterminable par télédétection classique sans intervention de données LIDAR.

Niveau 3 : Informations complémentaires qui ne sont plus liées directement à la description physionomique du milieu. Il s'agit par exemple des pratiques agricoles que subissent les herbacées : **fauchée** ou **pâturée**. Une information à prendre en compte est l'hydromorphie c'est-à-dire la teneur en eau du sol.

4. SEGMENTATION ADOPTÉE

La segmentation doit représenter le cadre de référence spatiale à la fois pour le télédécteur et pour l'expert botanique en charge des relevés de terrain. Par conséquent, les deux acteurs doivent s'accorder sur le niveau de segmentation le plus en concordance avec l'échelle de précision et le niveau de nomenclature désiré.

Dans le cadre du projet Carhab et pour les milieux ouverts de basse altitude, c'est la segmentation brute de l'IGN qui a été retenue. Cette couche est issue de la segmentation multirésolution des orthophotographies Infra-rouge couleur (IRC, à 0,5 m de résolution) de l'IGN réalisée à travers le logiciel PYRAM. C'est le niveau 2 qui sert de référentiel pour la cartographie des végétations naturelles et semi naturelles. Dans le cas des milieux ouverts d'altitude, la mosaïque paysagère en présence impose une segmentation plus fine que celles proposées par l'IGN.

La segmentation est exploitée sous un format compatible avec un Système d'informations Géographiques (SIG), généralement en format de fichier de forme (shapefile).

5. PRODUIT ATTENDU

A la suite de ce chapitre, nous présentons les démarches spécifiques à suivre lorsqu'il s'agit de valoriser une carte de végétation préexistante (6) ou de créer de nouveaux échantillons (7). Cependant, le produit final attendu doit se présenter sous la même forme.

Les polygones sites échantillons doivent avoir les caractéristiques suivantes :

- Ils doivent être suffisamment homogènes. Pour ce faire, ils se basent sur la segmentation de référence IGN du niveau 2.
- Ils respectent une nomenclature préétablie. C'est la nomenclature physionomique du niveau 2 et 3 qui doit être utilisée (se référer à la partie 3).
- Ils sont suffisamment représentatifs statistiquement pour permettre un bon apprentissage des algorithmes de classification. Idéalement au moins 20 échantillons par types nomenclatureaux existants doivent être générés.
- Ils doivent être bien répartis dans l'espace de la zone d'étude pour représenter la variabilité spatiale de la zone. Par conséquent, les polygones strictement voisins sont à éviter.

- Ils doivent respecter l'échelle d'étude. Les polygones devront avoir une surface supérieure à 0,5 hectares.

Le produit final doit se présenter sous la forme suivante :

- Couche SIG en shapefile (.shp)
- Couche projetée en Lambert 93
- Nom du fichier : `echant_physio_MOBA_AAMMJJ_zone.shp`
 - *AAMMJJ* : date de la production. AA : Année, MM Mois, JJ Jours. Ex 130415 pour le 15 avril 20013.
 - *zone* : site test ou numéro du département
- Couche polygonale issue du fichier segmentation renseigné dans une table attributaire structurée avec les « *CHAMPS* » suivants :
 - « *ID_SEGMENT* » identifiant unique numérique du segment
 - « *PHYSIO_NIVEAU1* » : *qui prend les valeurs suivantes* :
 - **Eau ; Minéral ; Herbacée ; Ligneux bas ; Ligneux hauts**
 - « *PHYSIO_NIVEAU2* » qui prend les valeurs suivantes
 - **Eboulis ; sol nu ; roche, herbacé - minéral**
 - la productivité des herbacées : **herbacée à phytomasse faible ; herbacée à phytomasse moyenne ; herbacée à phytomasse forte**
 - **ligneux bas mixte ; ligneux bas dense ; ligneux bas rocailleux ; fourré – fructifiée**
 - « *PRATIQUE* », spécifique aux milieux ouverts de basse altitude qui se décline en deux valeurs : **fauchée ; pâturée**
 - « *HYDROMORPHIE* » se décline en : **hydromorphe ; mésomorphe ; xéromorphe**

6. VALORISATION DES CARTES EXISTANTES :

Cette partie sur la valorisation des cartes de végétation préexistantes concerne uniquement les cartes déjà intégrées dans un SIG sous un format vectoriel. Cette méthode semblerait *a priori* plus rapide puisqu'il s'agit de mobiliser une carte de végétation déjà existante pour créer des sites d'entraînement compatibles avec l'étude en cours. Cependant, cette méthode nécessite plus de technicité en termes d'opérations géomatiques et de connaissance experte sur les nomenclatures utilisées en cartographies végétales. De plus, cette méthode nécessite une phase de vérification car le document peut être désuet ou non approprié. La carte doit présenter des compatibilités typologiques, spatiales et temporelles suivantes :

- compatibilité typologique : mettre en adéquation la typologie des objets avec la typologie physionomique définie en amont (traduire des codes Corine Biotoques en physionomie).
- compatibilité d'échelle : le document cartographique exploité doit être d'un niveau d'échelle supérieure ou égale au 1 :25 000 (par ex. : 1 :5000 ; 1 :10 000 ; 1 : 20 000 ...)
- compatibilité géométrique : la donnée doit être polygonale et non ponctuelle ou linéaire. La carte doit être facilement transférable dans le système de projection de la carte de segmentation.
- compatibilité temporelle : la carte doit être en cohérence temporelle avec l'imagerie satellitaire supportant la classification au risque de se retrouver avec des changements d'occupation du sol.

Après avoir contrôlé ces différentes compatibilités, l'opérateur peut suivre la démarche suivante :

1) **Nettoyage de la carte d'origine.** Pour ce faire, il faut préalablement sélectionner les habitats considérés comme non-complexes et de taille supérieure à 0,5 hectare.

2) **Transformation de la nomenclature d'origine vers la nouvelle nomenclature physionomique.** Cette étape est primordiale et nécessite une bonne connaissance des nomenclatures utilisées pour la cartographie des végétations. Il s'agit de retraduire la nomenclature de la carte au travers d'une table de correspondance selon les champs décrit en partie 5.

3) **Transformation géométrique éventuelle.** Les objets voisins appartenant à la même nouvelle classe sont fusionnés. De plus, la carte de végétation d'origine, peut subir une projection semblable à celle de la segmentation IGN c'est-à-dire en Lambert 93.

3) **Transfert vers la segmentation de référence.** Les descriptions retraduites des polygones de la carte d'origine doivent être retranscrites sur les polygones de la segmentation IGN. Cette opération consiste à réaliser une jointure spatiale dans un SIG.

4) **Contrôle visuel du transfert.** Le transfert vers la segmentation doit être absolument contrôlé à posteriori. Il s'agit d'une part de contrôler le passage d'une géométrie et de s'assurer que les types physionomiques sont encore d'actualité. Pour cette dernière étape une validation par photo-interprétation doit être réalisée en utilisant les fonds orthophotographiques de l'IGN qui ont été utilisés pour la segmentation.

5) **Contrôle de terrain.** Une campagne de terrain peut être envisagée si nécessaire afin de compléter l'information notamment sur la phytomasse des herbacés, les pratiques anthropiques (pâturage, fauche) ou l'hydromorphie.

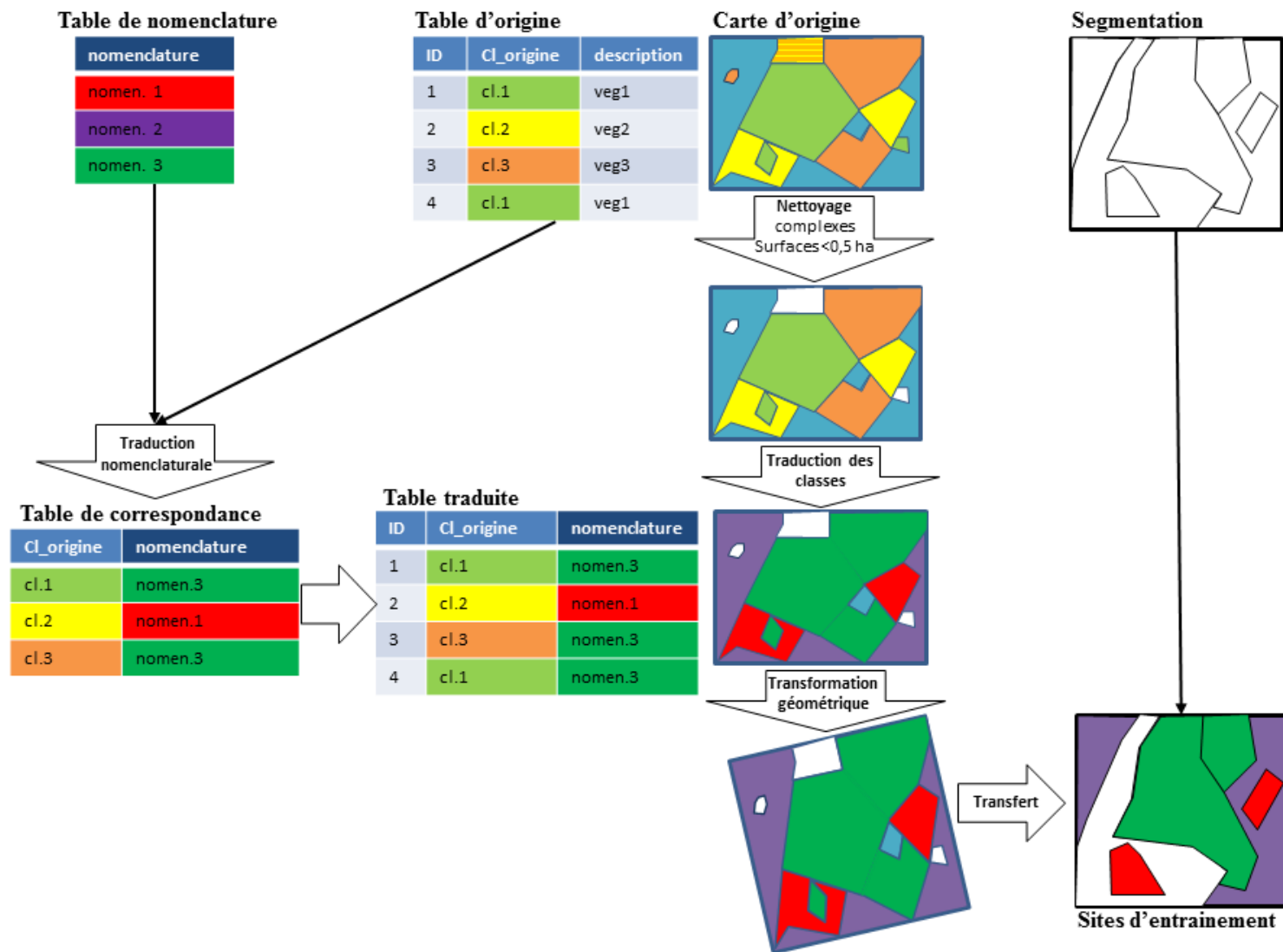


Figure 3. Etape de valorisation de carte existante

7. CRÉATION DE NOUVEAUX ECHANTILLONS

Un échantillonnage bien conçu spécifiquement dans le but de créer des sites d'entraînement pour la télédétection est certainement plus efficace que toute récupération de description de terrain qui n'était pas destinée à cet objectif. Cette méthode est plus simple à mettre en place techniquement.

1) **Visualisation dans un SIG.** La segmentation vectorielle, comportant les polygones qui vont être sélectionnés, est superposée à l'orthophotographie de l'IGN. La couche de segmentation doit comporter une table attributaire avec les champs décrit en partie 5.

2) **Choix de sites d'entraînement sur la base d'une photo-interprétation.** Au moins 20 polygones par classe nomenclaturale sont sélectionnés et renseignés *a priori* dans la table attributaire.

3) **Contrôle de terrain.** Une campagne de terrain doit être associée à la création des sites d'entraînement notamment afin de renseigner sur la productivité des herbacés, les pratiques anthropiques (pâturage, fauche) et l'hydromorphie.

Le couplage de cette méthode avec l'utilisation de cartes existantes est également une solution pour créer de nouveau sites échantillons.

8. RÉFÉRENCES

- Girard, M.-C., and Girard, C.M. (2010). *Traitement des données de télédétection environnement et ressources naturelles* (Paris: Dunod).
- Lang, S., Schöpfer, E., and Langanke, T. (2009). Combined object-based classification and manual interpretation—synergies for a quantitative assessment of parcels and biotopes. *Geocarto International* 24, 99–114.
- Nagendra, H. (2001). Using remote sensing to assess biodiversity. *International Journal of Remote Sensing* 22, 2377–2400.
- Weinke, E., Lang, S., and Preiner, M. (2008). Strategies for semi-automated habitat delineation and spatial change assessment in an Alpine environment. In *Object-Based Image Analysis*, T. Blaschke, S. Lang, and G. Hay, eds. (Springer Berlin Heidelberg), pp. 711–732.