

Chapitre 14. Modéliser des données paysagères pour les intégrer la mosaïque paysagère dans les modèles de répartition des espèces sentinelles du climat

Auteurs : Emmeline Lobry, Contributeurs : Laurent Couderchet, Fanny Mallard

Résumé : Les modèles corrélatifs visant à étudier le lien entre l'occurrence d'une espèce et des variables environnementales et climatiques, doivent prendre en compte la mosaïque paysagère. Or, l'intégration de variables paysagères dans ce type de modèle impose des contraintes techniques : elles doivent être spatialisées, selon une projection et une emprise géographique similaires à celles des autres variables explicatives utilisées dans le modèle. La conformité des résolutions spatiales et temporelles pose des questions à la fois théoriques et méthodologiques. En effet, les données paysagères sont fréquemment créées avec des bases de données d'occupation du sol, à des résolutions spatiales bien plus fines que celle des données climatiques. Ce chapitre vise à sélectionner les indices paysagers les plus pertinents vis-à-vis de chaque espèce et teste une méthode pour les combiner en un raster unique à intégrer dans le modèle de répartition d'espèce (SDM).

Mots-clés : données, *Hyla meridionalis*, *Hyla molleri*, indices paysagers, occupation du sol, présence, résolution.

14.1 Introduction

Les modèles de répartition d'espèces se focalisent généralement sur la corrélation entre le climat et la présence de l'espèce. Si ce choix apparaît pertinent pour la compréhension à des échelons continentaux ou mondiaux, d'autres variables explicatives sont indispensables à l'étude de la répartition des espèces à des échelons plus fins. En effet, il est alors nécessaire d'intégrer des variables qualifiant l'habitat, le paysage, la végétation ou les interactions biotiques (FRANCESCO FICETOLA *et al.*, 2014). En négligeant ces variables, notamment la connectivité paysagère, la plupart des projections basées sur des modèles de niche écologique sont limitées (MESTRE *et al.*, 2017). Elles apparaissent fondamentales et devraient donc être explicitement incluses dans les modèles de répartition d'espèce visant à prédire la réponse des espèces face aux changements environnementaux (VALLECILLO *et al.*, 2009).

Pourtant, ces processus sont étudiés, notamment en écologie du paysage. La recherche en écologie du paysage a développé des indices mathématiques, les métriques paysagères, qui peuvent caractériser la composition et la structure du paysage et être mises en lien avec les processus écologiques à l'œuvre dans cet espace. Les indices calculés permettent d'exprimer certaines dynamiques, par exemple en renseignant sur la connectivité entre différents patchs d'habitat, ou sur leur fragmentation (MCGARIGAL & MARKS, 1995). Ces données sont importantes pour comprendre la capacité des espèces à se déplacer et à survivre, en lien avec le déplacement possible de leur aire de répartition sous l'effet du changement climatique (VALLECILLO *et al.*, 2009 ; MESTRE *et al.*, 2017). Ce type de modélisation ne vise pas la projection cartographique de la niche de l'espèce étudiée, mais en fournit toutefois elle aussi une caractérisation.

Ces deux approches peuvent être réunies : la caractérisation des habitats, souvent utilisée comme un indicateur de biodiversité, peut être utilisée comme variable(s) explicative(s) dans un SDM (BRADLEY & FLEISHMAN, 2008 ; VALLECILLO *et al.*, 2009 ; TATTONI *et al.*, 2012 ; WILSON *et al.*, 2013). Tenir compte de ces facteurs permettrait d'approcher les interrelations entre nature et société, et ainsi de dissocier les effets liés à l'anthropisation, de ceux liés au changement climatique. Cette hypothèse de recherche a été initiée en 2016 par F. Mallard (Cistude Nature) dans la mise en place de la démarche de recherche du programme les sentinelles du climat (MALLARD, 2016a). Une méthode alors proposée est de s'appuyer sur des indices paysagers permettant de replacer les milieux étudiés dans l'espace en étudiant leur taux de dégradation et les échanges entre les systèmes voisins dans le paysage (MALLARD & COUDERCHET, 2019). Ensuite, la rencontre en 2017 entre Cistude Nature et l'UMR PASSAGES – CNRS (L. Couderchet) a permis de développer une recherche exploratoire sur la prise en compte des éléments paysagers dans la modélisation des répartitions des espèces sentinelles du climat, menée par E. Lobry dans le cadre d'une thèse en géographie (co-encadrement : L. Couderchet UMR PASSAGES–CNRS et F. Mallard, Cistude Nature/UMR PASSAGES–CNRS).

Les métriques paysagères sont fréquemment calculées à partir de cartes d'occupation du sol. Or, ces cartes sont souvent produites pour répondre à plusieurs enjeux en matière d'aménagement du territoire. Ainsi, elles ne sont souvent pas assez précises et négligent les caractéristiques écologiques, qui sont fondamentales pour la faune et la flore (TATTONI *et al.*, 2012 ; FRANCESCO FICETOLA *et al.*, 2014). Elles sont généralement réalisées au format vectoriel et proposent dès lors une visualisation discrète : les espaces sont délimités et définis par des catégories, par opposition à des cartes au format raster, réalisées par télédétection, où la donnée peut être continue, la seule limite visible étant le périmètre de la cellule (COUDERCHET, 2008). Cette discrétisation en types d'occupation du sol peut sembler inadaptée à l'objectif poursuivi. Cependant, cette problématique peut être nuancée dans le cas où l'étude porte sur une espèce spécialiste d'un habitat ainsi catégorisé (FRANCESCO FICETOLA *et al.*, 2014).

De plus, ces bases de données sont généralement disponibles en France, ce qui permet une certaine reproductibilité de la méthode proposée, et la compréhension des changements d'occupation du sol semble primordiale dans la mesure où ils agissent sur l'état de la biodiversité et *a fortiori* dans le contexte du changement climatique (VALLECILLO *et al.*, 2009).

Des variables de structure et de composition de la végétation peuvent aussi être obtenues par le traitement d'images, notamment la télédétection : images satellite multi-capteurs, LiDAR, drone, apportent des données sur la présence de la végétation (par exemple, indice de végétation par différence normalisé NDVI) ou sa structure verticale (hauteurs de végétation) et plusieurs auteurs ont ainsi testé l'utilisation de ces données dans des SDM (BRADLEY & FLEISHMAN, 2008 ; VALLECILLO *et al.*, 2009 ; WILSON *et al.*, 2013 ; FRANCESCO FICETOLA *et al.*, 2014). Dans le cadre du programme, un stage réalisé en 2019 pour expérimenter une cartographie fine des habitats présents sur les sites de suivi à l'aide d'un drone (PREVITALI, 2019). La résolution spatiale est alors de l'ordre centimétrique et permet de renseigner sur la structure et la dynamique de la végétation : le Modèle Numérique d'Élévation, en mesurant la hauteur, amène à discriminer les strates herbacées, arbustives et arborées. Ainsi, dans le milieu des pelouses calcicoles, il a été possible de différencier deux sites aux dynamiques différentes : l'un étant entretenu, l'autre en cours de fermeture. Afin de réaliser ces cartographies par télédétection, l'adéquation entre la résolution spatiale et l'objectif de description du site de suivi a également été questionnée : les images source ayant une résolution centimétrique, la reconnaissance des milieux a nécessité un rééchantillonnage vers une résolution d'ordre décimétrique.

Des questions théoriques et méthodologiques se posent avec l'utilisation des métriques paysagères ou de l'occupation du sol en tant que variables explicatives dans un SDM. En effet, la programmation du SDM, avec le package BIOMOD2 développé en langage R, impose l'emploi d'un lot de variables explicatives spatialisées, construites au format raster selon une projection, une emprise géographique et des résolutions spatiale et temporelle similaires (GEORGES & THUILLER, 2013). Cette contrainte interroge la pertinence d'utiliser, pour construire les données paysagères, les résolutions spatiale et temporelle des données climatiques, alors que les phénomènes ne se déroulent pas nécessairement dans les mêmes échelles d'espace et de temps. Ainsi, pour correspondre à la résolution spatiale des variables climatiques, les métriques paysagères doivent être dégradées ou reconfigurées. D'une certaine façon, des résolutions spatiales différentes peuvent ainsi être mobilisées, toutefois leur concordance et leur articulation doivent être explicitées.

Pour combiner les données collectées à l'échelle du site et l'échelle régionale du rendu du programme, il est nécessaire de réfléchir à l'articulation de la résolution spatiale des données climatiques (de l'ordre du kilomètre), celle des données d'occupation du sol (de l'ordre de la centaine de mètres) et celle d'occurrence des espèces (points). Dans un précédent tome (MALLARD, 2019), l'importance de cette question de la résolution spatiale comme un préalable a été démontrée. D'une part, parce que la résolution spatiale de création des données d'occupation du sol, et donc du paysage, a des conséquences sur la manière dont sont reconnues les entités qui les composent. D'autre part, la résolution de travail permet, ou non, d'obtenir des informations pertinentes pour l'objectif de l'étude. En effet, la résolution spatiale peut constituer un facteur de compréhension : l'analyse des relations entre les données paysagères de diverses résolutions spatiales et l'occurrence des espèces permettrait d'indiquer la résolution spatiale permettant de reconnaître certains arrangements spatiaux, par exemple la connectivité paysagère.

L'objectif de ce chapitre est d'appliquer la méthode développée précédemment (Lobry, 2020 in Mallard, 2020) en différenciant les deux espèces *Hyla molleri* et *Hyla meridionalis*. Il s'agit d'identifier les indices paysagers les plus pertinents et de les combiner en un indice synthétique qui puisse être intégré au modèle de répartition d'espèces. Toutefois, cette synthèse ne doit pas s'effectuer au

détriment de la compréhension du sens de cet indicateur et la traçabilité de sa construction est donc mise en avant.

14.2 Matériel et Méthodes

14.2.1 Données utilisées

Les données d'occurrence des deux espèces de rainettes étudiées *Hyla meridionalis* et *Hyla molleri* sont issues de la base de données répertoriées par l'association Cistude Nature depuis 2004 dans le cadre notamment des Atlas des Amphibiens et Reptiles d'Aquitaine (BERRONEAU, 2014a). Ces données prennent la forme d'un fichier vectoriel de points.

Des indices paysagers ont été calculés d'après deux cartes d'occupation du sol issues de deux bases de données : CORINE Land Cover (2018) et le Référentiel néo-aquitain d'occupation du sol à grande échelle (2015). Ces deux bases de données ont été construites à des résolutions différentes : les unités minimales cartographiées mesurent respectivement 25 ha et 1 ha (0.1 ha pour les classes de « territoires artificialisés »). Deux cartes ont été créées et décrivent le type d'occupation du sol présent majoritairement dans des mailles de 500 m et 100 m respectivement (Lobry, 2019 in (MALLARD, 2019)).

Les métriques paysagères ont été calculées avec le logiciel Fragstats sous la forme spatialisée avec l'outil de calcul dans des fenêtres coulissantes. Des paramètres liés à l'épaisseur de lisière, le contraste et la similarité ont été définis pour le calcul des indices concernés (Annexe 33). Plusieurs tailles de fenêtres coulissantes et les deux cartes d'occupation du sol ont servi à calculer les indices afin de comparer leur relation à la donnée de présence des rainettes (Lobry, 2020 in MALLARD, 2020).

L'analyse a été effectuée dans l'emprise du département de Gironde. Ce choix correspond à la disponibilité des données du Référentiel néo-aquitain d'occupation du sol à grande échelle (OCS-GE) par département et à la possibilité future de réfléchir à l'articulation de ce travail avec les données paysagères récoltées sur quelques sites girondins lors de travaux précédents (LOBRY *et al.*, 2018, PREVITALI, 2019).

14.2.2 Analyses statistiques et d'images

Précédemment, les analyses tenaient compte des points de présence de l'une ou l'autre des espèces indifféremment. Dans ce chapitre, la variable à expliquer s'est complexifiée puisqu'elle est passée d'une variable booléenne (présence ou pseudo-absence quelle que soit l'espèce), à une variable ayant trois valeurs possibles (*Hyla molleri*, *Hyla meridionalis* ou pseudo-absence).

- **Analyse par arbre(s) de décision**

L'analyse statistique de ces données reprend la méthode développée précédemment (LOBRY, 2020). Elle repose sur la méthode de l'arbre de décision. L'outil BaobARD est une interface développée par Grégoire Le Campion (UMR PASSAGES) qui propose la construction d'un arbre de décision en combinant plusieurs fonctions du logiciel R. Cet outil est disponible en ligne (https://analytics.humanum.fr/Gregoire.LeCampion/Arbre_de_decision/).

Chaque arbre de décision sélectionne les indices paysagers les plus explicatifs et précise la valeur de cet indice qui permet de différencier les deux sous-branches créées, avec la plus grande homogénéité intra-classe et la plus grande hétérogénéité inter-classe possibles.

Trois arbres de décision ont été réalisés. Le premier prend en compte l'ensemble des rasters d'indices calculés, quelles que soient la base de données d'occupation du sol utilisée ou la taille de fenêtre coulissante paramétrée. Le second tient compte des rasters d'indices calculés d'après CORINE Land Cover (CLC), à toutes les tailles de fenêtre coulissante calculées, soit de 1.5 km à 5.5 km. Enfin, le troisième analyse les indices calculés d'après le Référentiel néo-aquitain d'occupation du sol pour l'ensemble des tailles de fenêtre coulissante, soit de 300 m à 5.5 km.

L'analyse de ces indices identifiés comme prédicteurs nous permettent d'observer les caractéristiques paysagères qui paraissent importante pour l'une, l'autre, ou les deux espèces étudiées. Outre cette première analyse descriptive, les indices sélectionnés constituent une base pour la suite de la méthode de traitement visant à tendre vers une sélection d'indices suffisamment peu nombreux pour être intégrés dans le modèle de répartition global (SDM).

- *Analyses d'images*

Une méthode a été testée pour combiner les rasters d'indices paysagers et en tirer une information synthétique, tout en conservant une certaine traçabilité du sens écologique de ces indices. Elle s'appuie sur des méthodes d'analyse d'images telle que pratiquée fréquemment en télédétection.

D'abord, une Analyse en Composantes Principales (ACP) a été réalisée avec les rasters d'indices paysagers sélectionnés dans chacun des trois arbres de décision. Les rasters issus de ce pré-traitement, correspondant aux trois premiers axes de l'ACP, ont été utilisés comme rasters d'entrée dans une classification par K-moyennes, paramétrée arbitrairement pour obtenir 10 classes au maximum.

Le résultat de cette classification permet de dessiner des espaces dont les caractéristiques paysagères sont relativement homogènes. L'analyse visuelle et leur description amène une compréhension du territoire girondin différente de celle apportée par les bases de données d'occupation du sol. L'analyse des « signatures spectrales » permet de conserver les informations contenues dans les indices paysagers : nom de l'indice, valeurs minimale, moyenne et maximale de la classe.

Enfin, une jointure spatiale réalisée avec le logiciel QGIS permet de chercher le lien entre les points de présence des deux espèces de rainettes pour analyser en particulier les classes où se trouvent les espèces.

L'ensemble de cette méthode d'analyse d'images est présenté dans le schéma suivant (Fig. 271).

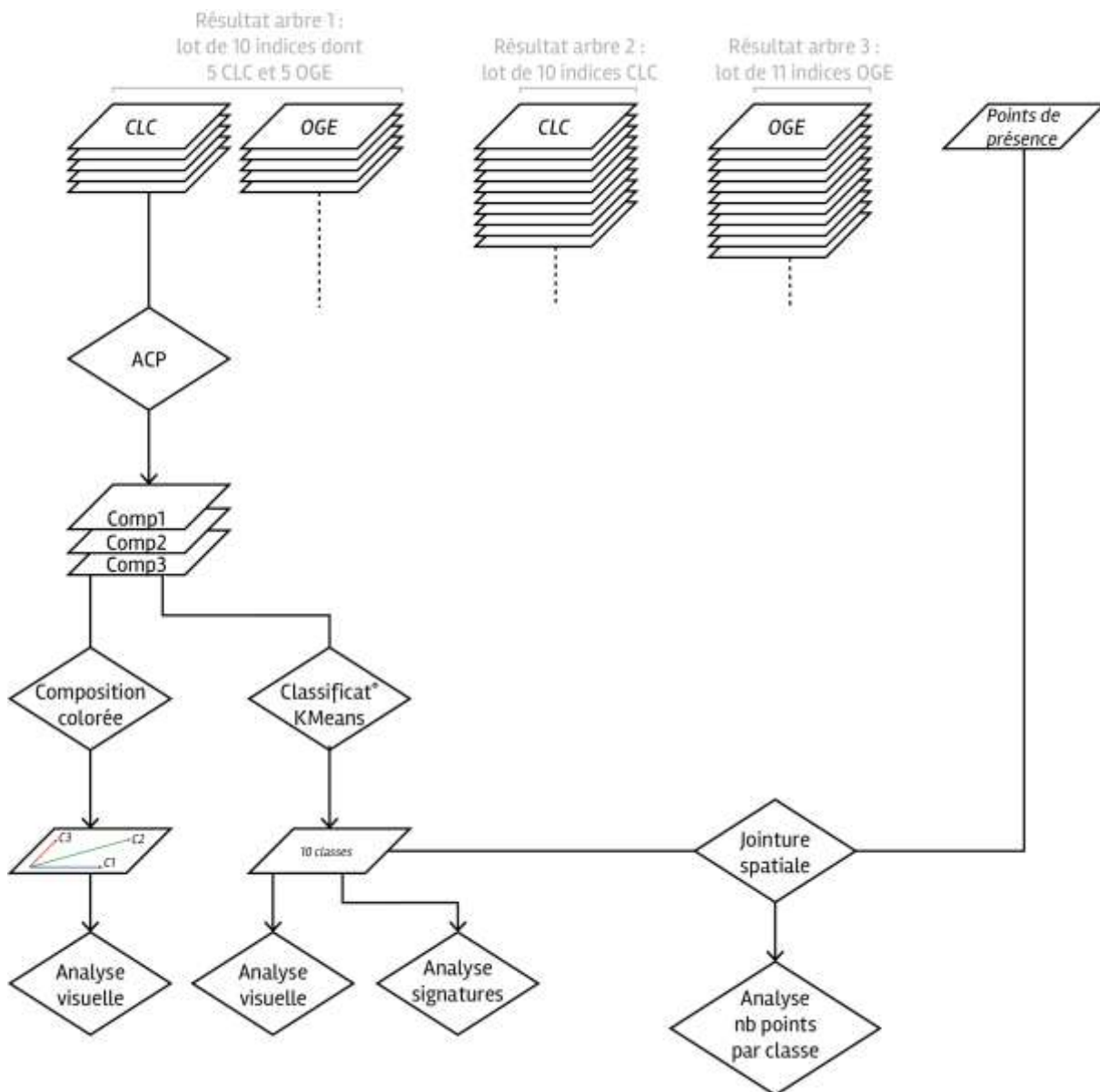


Fig. 271. Étapes de traitement des images d'indices paysagers

14.3 Résultats

Les résultats présentés dans ce chapitre sont organisés en deux parties. La première vise à identifier un lot d'indices paysagers pertinents pour chacune des espèces de rainette par une analyse statistique avec la méthode de l'arbre de décision. La seconde montre les résultats préliminaires d'une méthode, basée sur l'analyse d'images, pour combiner les indices paysagers en un raster unique afin de l'intégrer au modèle de répartition global.

14.3.1 Analyses statistiques par arbres de décision : identifier des indices paysagers pertinents

- *Analyse simultanée de toutes les variables, quelle que soit la base de données d'occupation du sol ou la taille de fenêtre coulissante*

L'arbre de décision présenté ci-dessous (Fig. 272) réalisé a pris en compte l'ensemble des variables paysagères. Il montre une classification des 2236 points de présence et pseudo-absence en 13 classes,

dont 5 composées majoritairement de points de présence de *Hyla meridionalis* (74% en moyenne) et 2 composées majoritairement de points de présence de *Hyla molleri* (72% en moyenne). Dans les classes des *Hyla meridionalis*, aucune co-présence des deux espèces n'apparaît. En revanche, dans les 2 classes de *Hyla molleri*, les deux espèces sont co-présentes (23% de *Hyla meridionalis* en moyenne dans ces classes).

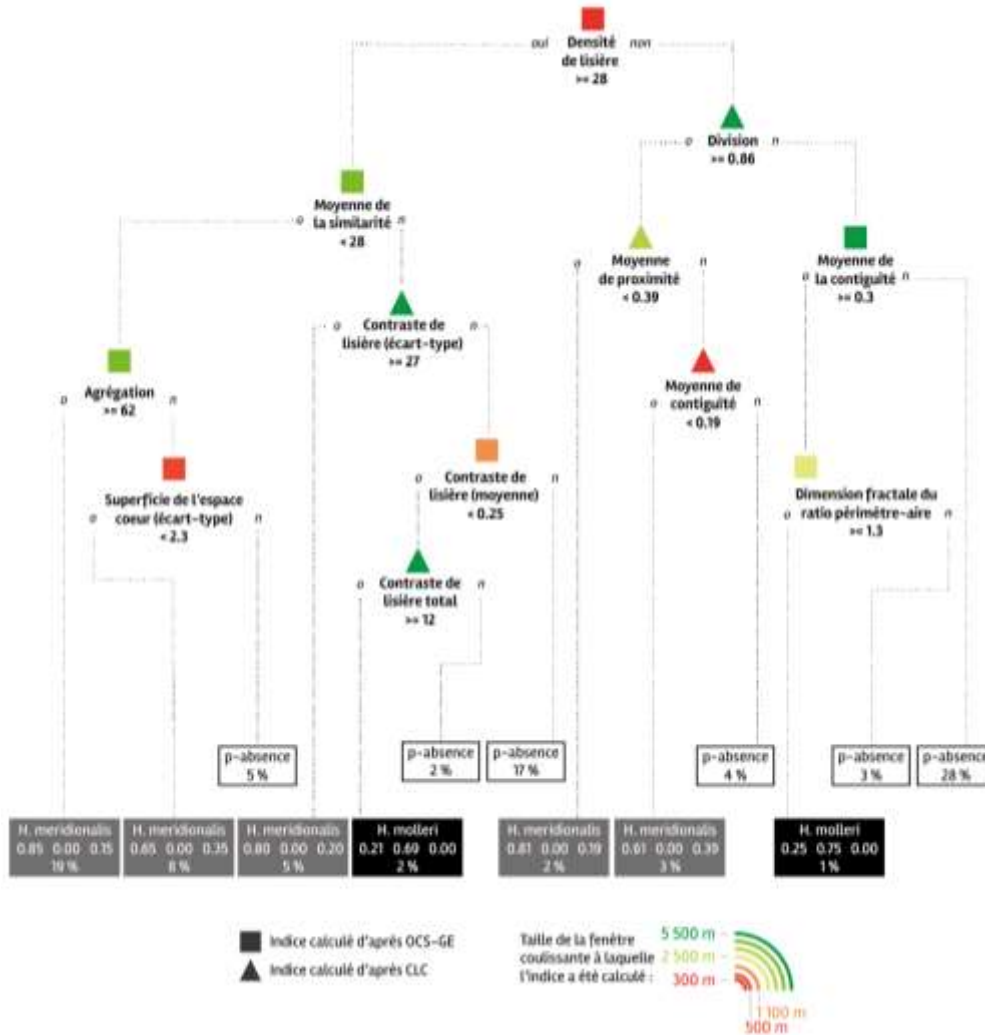


Fig. 272. Deux classes de *Hyla molleri* et cinq classes de *Hyla meridionalis* dans l'arbre de décision prenant en compte l'ensemble des indices paysagers

La figure suivante (Fig. 273) représente les indices présents dans cet arbre, c'est-à-dire ceux qui expliquent le mieux la présence de chacune des espèces de rainette en Gironde. Ils sont placés sur des cercles concentriques selon la taille de fenêtre coulissante à laquelle l'indice a été calculé. Ils sont également caractérisés selon leur mode de construction, c'est-à-dire la base de données d'occupation du sol utilisée : les triangles pour CORINE Land Cover et les carrés pour le Référentiel néoquitain d'occupation du sol.

Les indices prédicteurs de chaque espèce sont précisés sur chaque groupe de cercles. La majorité des indices peut être liée à la présence des deux espèces : densité de lisière, division, similarité moyenne, écart-type du contraste de lisière, contiguïté moyenne. Il faut noter toutefois une variation pour ce dernier indice, explicatif pour la présence de chacune des espèces : pour *Hyla molleri*, cet indice est explicatif lorsqu'il est calculé d'après le Référentiel néoquitain tandis que pour *Hyla meridionalis*,

il l'est lorsqu'il est basé sur CORINE Land Cover. Dans les deux cas, il est explicatif lorsqu'il est calculé dans une fenêtre coulissante de 5.5 km de côté.

Trois indices diffèrent entre les deux espèces. L'écart-type de la superficie de l'espace cœur, la moyenne de l'indice de proximité et l'indice d'agrégation n'apparaissent pas explicatifs de la présence de la rainette ibérique *Hyla molleri* contrairement à la rainette méridionale *Hyla meridionalis*. A l'inverse, la dimension fractale du ratio périmètre–aire, le contraste de lisière moyen et total sont explicatifs de la présence de la rainette ibérique mais pas de la rainette méridionale.

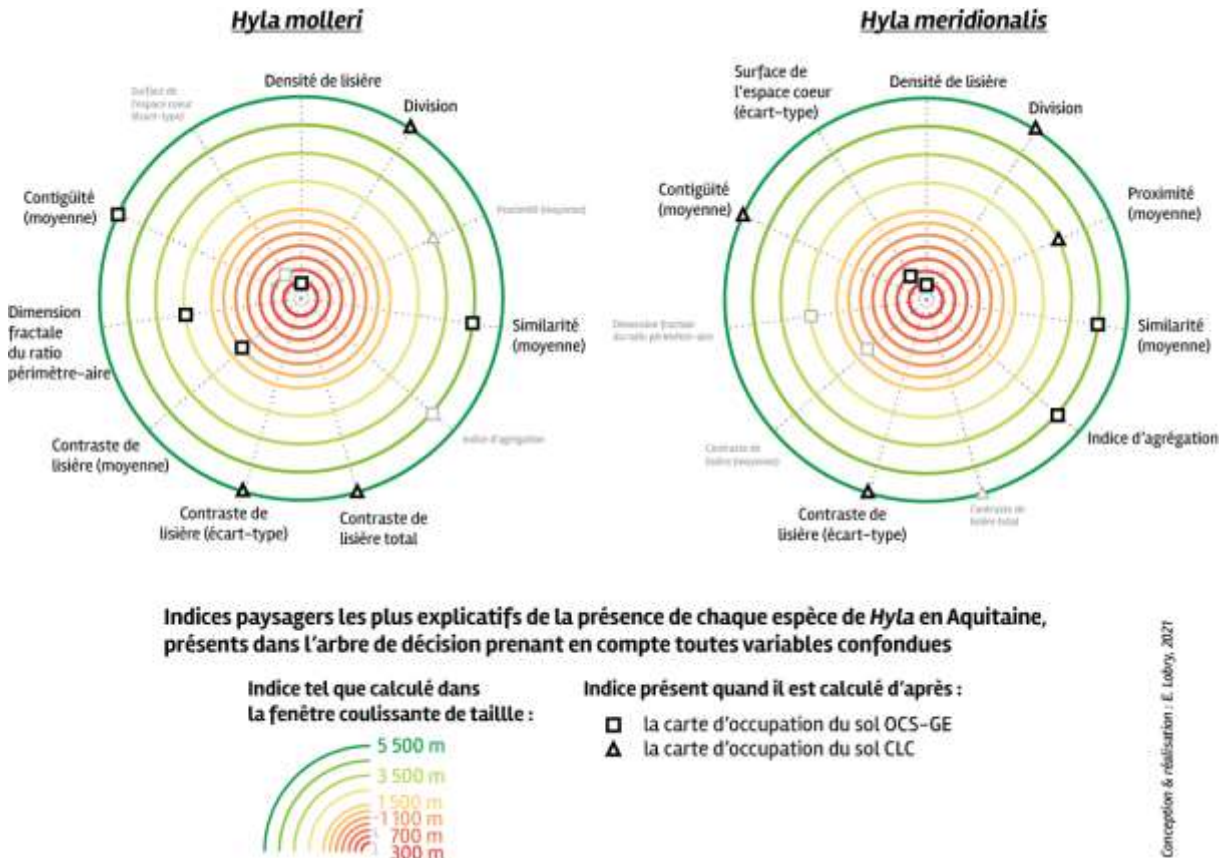


Fig. 273. La présence de *Hyla molleri* est principalement liée aux indices de lisière, celle de *Hyla meridionalis* plus en lien avec les « cœurs de biodiversité »

- **Analyse différenciée selon la base de données d'occupation du sol utilisée pour construire les indices**

Afin de pouvoir mener par la suite des analyses d'images, les rasters d'indices paysagers doivent être de même résolution. Ainsi, deux arbres de décision permettent de différencier les résultats obtenus avec chaque base de données d'occupation du sol.

Le premier arbre présenté ci-dessous (Fig. 274) correspond aux indices paysagers créés d'après les rasters construits avec CORINE Land Cover, de 500 m de résolution. Cet arbre aboutit à l'identification de 12 classes, dont 5 composées majoritairement de points de présence de rainette méridionale et une de rainette ibérique. Quelques points de présence de celle-ci se retrouvent dans des classes de rainettes méridionales (0.018% en moyenne). La classe des points de *Hyla molleri* est moins homogène que celles de *H. meridionalis* puisqu'ils ne représentent que 60% des points (73% en moyenne pour *H. meridionalis*) : les deux espèces sont donc co-présentes dans les espaces caractérisés par les indices paysagers identifiant ce groupe de points. En revanche, cette classe est définie par seulement 3 indices

paysagers tandis que les classes de points de *Hyla meridionalis* sont qualifiées par un nombre d'indices plus importants (4.6 en moyenne). Enfin, la classe des points de rainette ibérique est caractérisée par des indices calculés dans de grandes fenêtres coulissantes (4.5 et 5.5 km de côté, soit une superficie de 81 à 121 mailles à cette résolution) tandis que celles des points de rainette méridionale sont déterminées par des indices calculés dans des fenêtres de tailles plus variables.

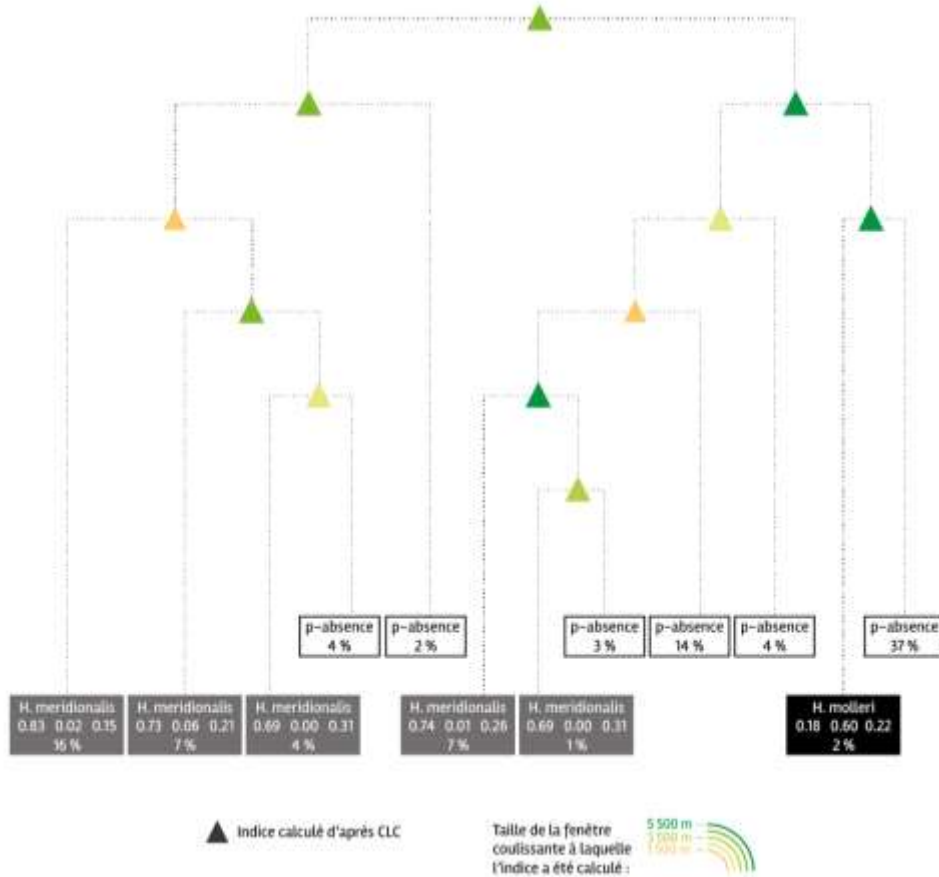


Fig. 274. Les indices paysagers construits d'après CORINE Land Cover plus appropriés pour l'espèce *Hyla meridionalis*?

L'arbre suivant (Fig. 275) correspond aux indices paysagers calculés d'après le Référentiel néo-aquitain d'occupation des sols (OCS-GE). Il identifie 12 classes dont 4 composées principalement de points de présence de *Hyla meridionalis* et 2 de *Hyla molleri*. Comme avec les indices construits d'après CLC, les classes composées principalement de points de présence de rainette ibérique sont plus hétérogènes que celles de rainette méridionale : 63% en moyenne sont de l'espèce majoritaire pour les classes de *H. molleri*, tandis que ce chiffre s'élève à 75% en moyenne pour celles de *H. meridionalis*. Au contraire des indices construits à partir de l'occupation du sol d'après CLC, le nombre d'indices nécessaire à la création d'une classe suffisamment homogène est de 3.25 en moyenne pour les classes de *H. meridionalis* et de 4.5 en moyenne pour les classes de *H. molleri*. Ce nombre d'indices plus élevé amène aussi une plus grande hétérogénéité de la taille de fenêtre coulissante pertinente : les superficies minimales et maximales sont utilisées pour les deux espèces, toutefois la présence d'indices calculés dans des fenêtres de taille moyenne (1.1 km, 2.5 km) est plus notable pour les classes de *H. molleri* tandis qu'ils correspondent aux extrêmes (300 et 500 m, 4.5 et 5.5 km) pour les classes de *H. meridionalis*.

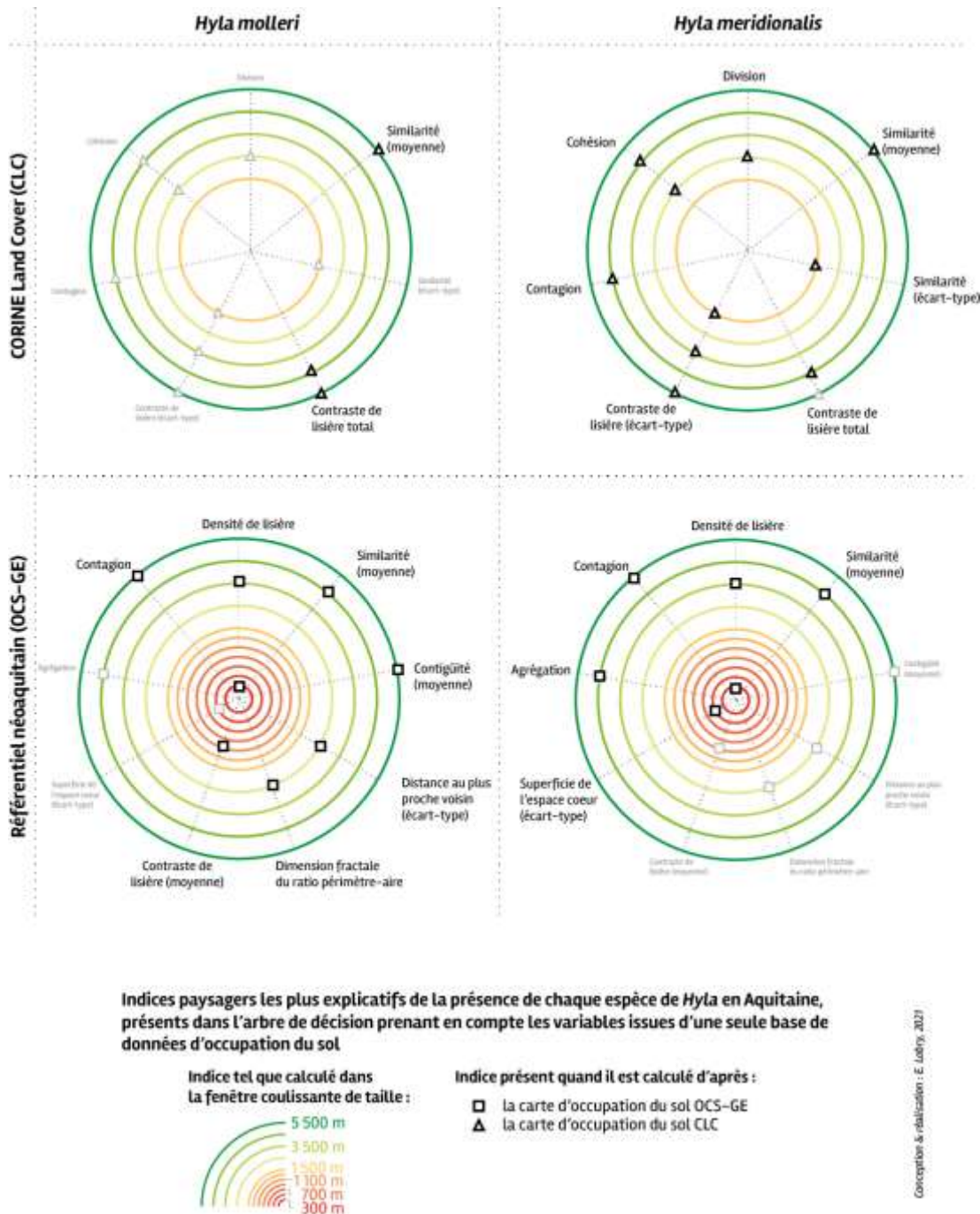


Fig. 276. Les indices de lisière sont de bons prédicteurs de la présence de *Hyla molleri* lorsqu'ils sont calculés d'après OCS-GE et de celle de *Hyla meridionalis* lorsqu'ils sont calculés d'après CLC

Cette première étape d'analyse statistique a permis d'identifier des lots d'indices paysagers pertinents d'une manière globale et selon une différenciation d'après les bases de données d'occupation du sol utilisées. Cette distinction est nécessaire pour l'analyse d'images puisque les lots de rasters traités doivent être de même résolution spatiale : 100 m pour les indices construits avec le Référentiel néo-aquitain, 500 m pour ceux construits d'après CORINE Land Cover. Le lot d'indices paysagers identifiés comme pertinents dans l'arbre de décision analysant simultanément l'ensemble des indices (Fig. 272) a été divisé en deux lots de 5 indices (cf. Fig. 271) en fonction de leur résolution spatiale. La partie suivante présentera les résultats obtenus par analyses d'images sur les 4 lots constitués.

14.3.2 Analyse d'images : combiner les rasters d'indices paysagers

Les premières analyses statistiques ont permis d'identifier quelques indices paysagers pertinents vis-à-vis des deux espèces étudiées. Pour réduire encore le nombre de rasters à intégrer dans le modèle de répartition d'espèces (SDM), une méthode d'analyse d'images a été testée. Un prétraitement a été réalisé pour chacun des quatre lots d'indices : une analyse en composantes principales (ACP) permet de combiner les informations majeures de chacune des images tout en gardant une traçabilité quant à la signification de chaque image résultante.

Pour chaque lot d'indices, les composantes principales sont exprimées d'une part de manière numérique et d'autre part dans de nouvelles images. Le tableau suivant (Tab. LXXVIII) présente le résultat obtenu pour l'ACP réalisée avec le lot de 5 indices construits à partir du Référentiel néo-aquitain (OCS-GE) identifiés dans le premier arbre de décision (cf. Fig. 272). Ce sont ces valeurs numériques qui permettent de suivre la contribution de chaque indice aux axes factoriels de l'ACP.

Les indices d'écart-type de la superficie de l'espace-cœur et la moyenne de l'indice de similarité constituent les deux extrémités de l'axe de la première composante. Les indices d'agrégation et la moyenne de similarité ceux de l'axe de la seconde composante. Enfin, les indices de densité de lisière et d'agrégation sont les extrémités de l'axe de la troisième composante. Ce sont donc ces indices qui sont principalement représentés dans les images des composantes déterminées par l'ACP.

Tab. LXXVIII. Les indices de similarité moyenne, l'écart-type de la superficie de l'espace-cœur, d'agrégation et de densité de lisière sont fortement discriminés dans les trois premières composantes

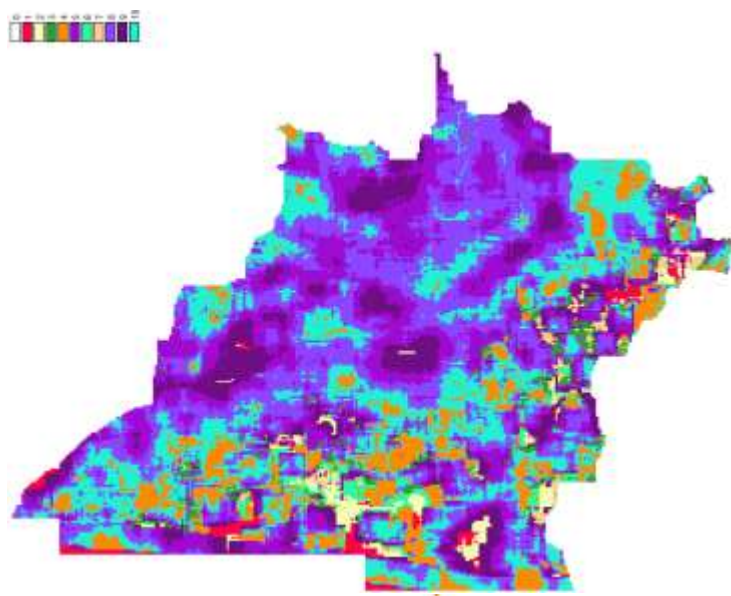
<i>Indice construit d'après OCS-GE</i>	<i>Taille de fenêtre coulissante</i>	CONTRIBUTION AUX COMPOSANTES PRINCIPALES		
		<i>Composante ACP 1</i>	<i>Composante ACP 2</i>	<i>Composante ACP 3</i>
Densité de lisière	300 m	13 741.49	-27 356.46	-32 896.89
Ecart-type de la superficie de l'espace-cœur	500 m	2 077.98	-1 985.31	-134.37
Moyenne du contraste de lisière	1 100 m	4 050.65	-6 724.54	-3 612.62
Indice d'agrégation	4 500 m	53 815.34	-53 153.70	1 571.16
Moyenne de l'indice de similarité	4 500 m	117 473.50	-1 744.82	-12 095.24

Le tableau suivant (Tab. LXXIX) présente le résultat obtenu pour l'ACP réalisée avec le lot de 5 indices construits à partir de CORINE Land Cover (CLC) identifiés dans le premier arbre de décision (cf. Fig. 272). Ici, c'est la moyenne de l'indice de contiguïté et l'écart-type du contraste de lisière qui contribuent le plus à la première composante. La deuxième composante renseigne principalement sur le contraste de lisière : écart-type et total. La troisième composante apporte une information différente, liée aux indices de proximité et de contiguïté. Globalement, ce sont l'effet de lisière et la notion de contiguïté qui sont exprimés dans les images des composantes déterminées par cette ACP.

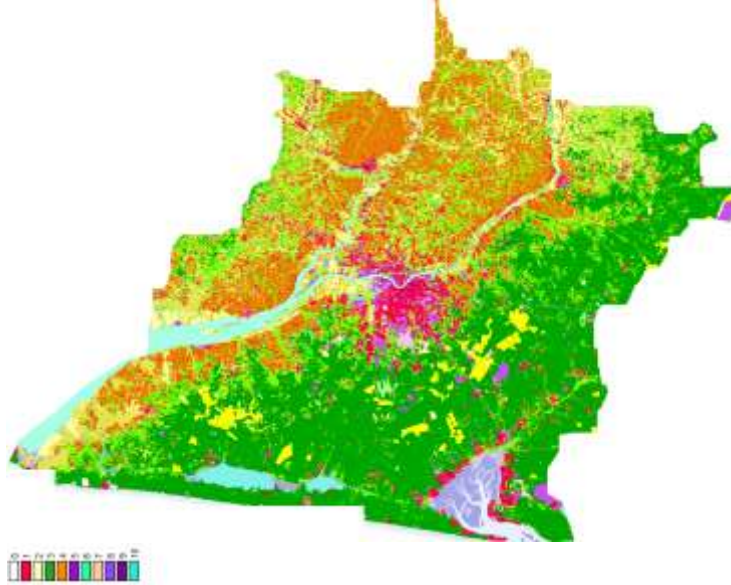
Tab. LXXIX. Les effets de lisière et la contiguïté contribuent le plus fortement aux premières composantes

<i>Indice construit d'après CLC</i>	<i>Taille de fenêtre coulissante</i>	CONTRIBUTION AUX COMPOSANTES PRINCIPALES		
		<i>Composante ACP 1</i>	<i>Composante ACP 2</i>	<i>Composante ACP 3</i>
Moyenne de l'indice de proximité	3 500 m	97.70	4.92	150.74
Moyenne de l'indice de contiguïté	5 500 m	0.13	0.07	-0.99
Division	5 500 m	154.30	8.91	43.63
Ecart-type du contraste de lisière	5 500 m	3 442.89	-225.11	62.73
Contraste de lisière total	5 500 m	2 011.72	611.33	63.03

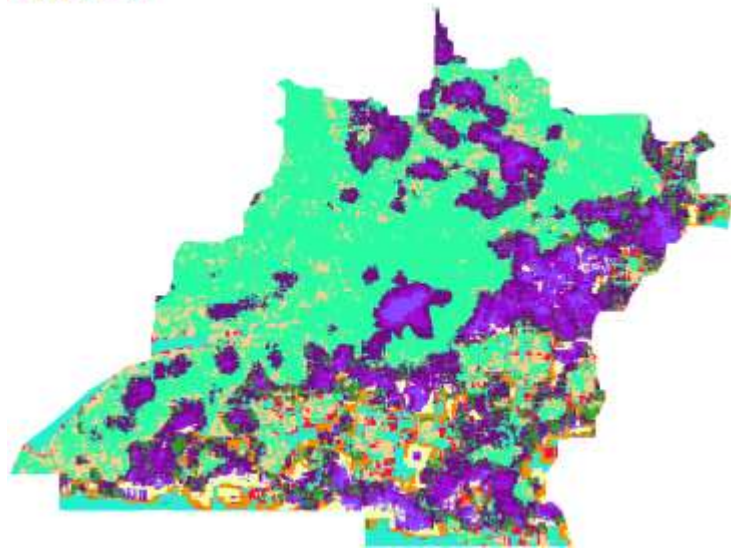
Pour chacun des quatre lots, les trois images des composantes principales ont été utilisées pour réaliser une classification non dirigée avec la méthode des K-moyennes. Les figures suivantes présentent les images de ces classifications, respectivement avec les indices construits d'après OCS-GE (Fig. 277) et d'après CLC (Fig. 278).



Classification par K-Moyennes des 3 composantes de l'ACP réalisée avec 5 indices OGE (résolution 100 m)

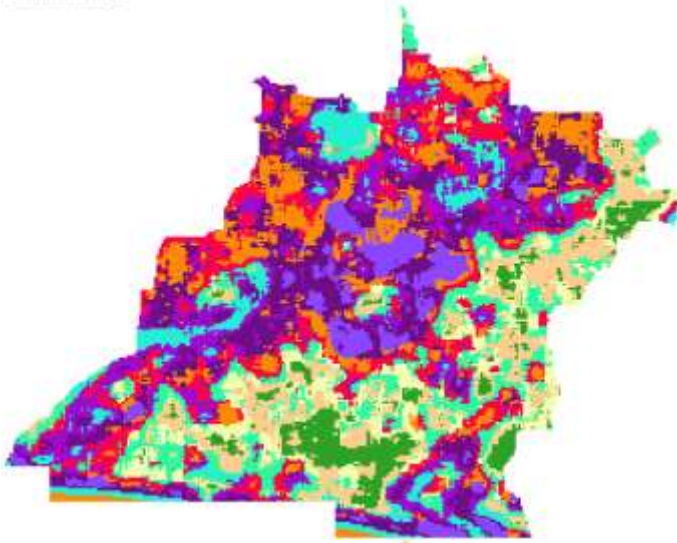


Types d'occupation du sol dominants dans une maille de 100 m d'après le Référentiel néoquittain

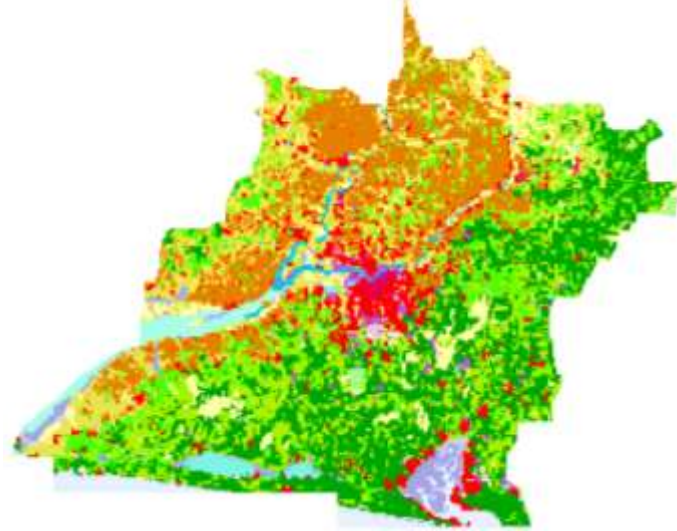


Classification par K-Moyennes des 10 composantes de l'ACP réalisée avec 10 indices OGE (résolution 100 m)

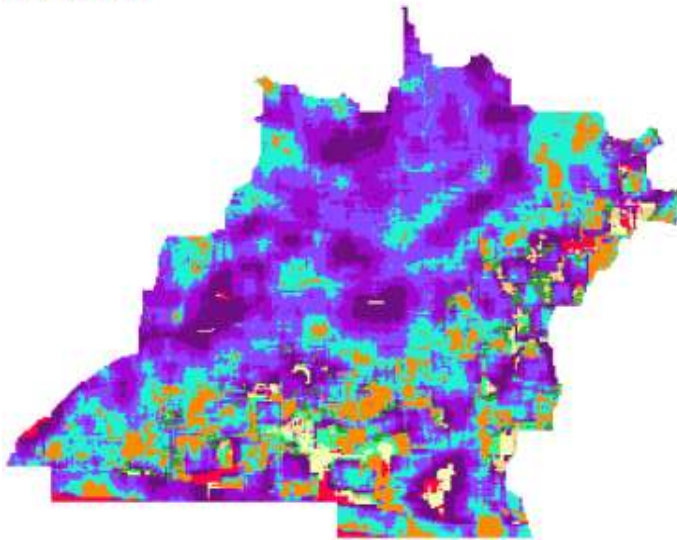
Fig. 277. Classifications réalisées à partir des indices construits d'après le Référentiel néoquittain



Classification par K-Moyennes des 3 composantes de l'ACP réalisée avec 5 indices CLC (résolution 500 m)



Types d'occupation du sol dominants dans une maille de 500 m d'après CORINE Land Cover



Classification par K-Moyennes des 3 composantes de l'ACP réalisée avec 11 indices CLC (résolution 500 m)

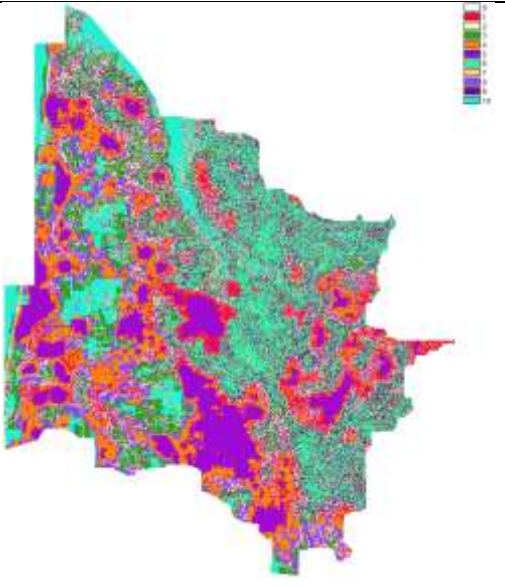
Fig. 278. Classifications réalisées à partir des indices construits d'après CORINE Land Cover

Une jointure spatiale entre ces rasters et le fichier de points de présence, menée avec un logiciel SIG, permet d'identifier les classes comprenant les plus grands effectifs de chaque espèce de rainette. Les tableaux suivants présentent ces résultats, respectivement pour les indices construits d'après OCS-GE (Tab. LXXX) et d'après CLC (Tab. LXXXI).

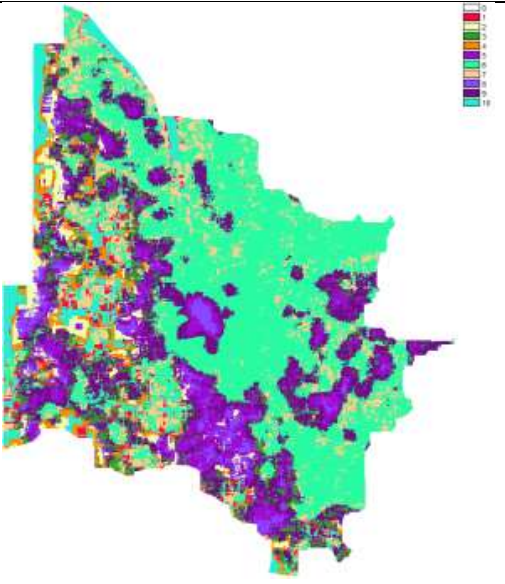
Dans l'emprise, il y a 8 points de présence de *Hyla molleri* et 456 points de *Hyla meridionalis*. Le nombre trop peu important de points de *Hyla molleri* ne permet pas d'analyser ces résultats. En revanche, la classification semble apporter une bonne discrimination des points de présence de rainette méridionale. Dans l'image classée avec les 5 indices OCS-GE, les classes 6 et 9 contiennent près de 77% de l'effectif des rainettes méridionales. Dans celle construite avec les 10 indices, la classe 6 en contient 80%. Ces classes correspondent à des habitats potentiellement favorables à cette espèce.

Tab. LXXX. Les classifications basées sur OCS-GE permettent de discriminer environ 80% de l'effectif de *Hyla meridionalis*

Classification d'après les 3 composantes de l'ACP réalisée sur les 5 indices OCS-GE (cf. Fig. 272)			
	Part des points dans les classes (%)		
	Toutes espèces	<i>Hyla meridionalis</i>	<i>Hyla molleri</i>
Classe 1	4.74	4.39	25.00
Classe 2	2.59	2.19	25.00
Classe 3	0.65	0.66	0.00
Classe 4	3.02	2.41	37.50
Classe 5	4.74	4.82	0.00
Classe 6	34.27	34.87	0.00
Classe 7	6.90	7.02	0.00
Classe 8	0.65	0.66	0.00
Classe 9	41.59	42.11	12.50
Classe 10	0.86	0.88	0.00






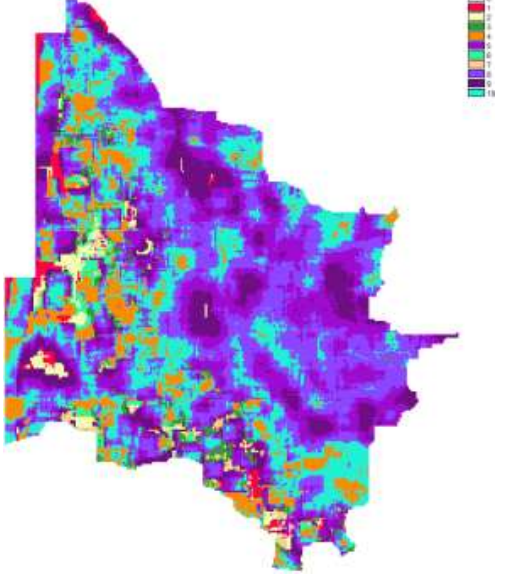
Classification d'après les 3 composantes de l'ACP réalisée sur les 10 indices OCS-GE (cf. Fig. 275 et Fig. 276)			
	Part des points dans les classes (%)		
	Toutes espèces	<i>Hyla meridionalis</i>	<i>Hyla molleri</i>
Classe 1	0.00	0.00	0.00
Classe 2	0.00	0.00	0.00
Classe 3	0.43	0.44	0.00
Classe 4	0.65	0.66	0.00
Classe 5	2.59	2.63	0.00
Classe 6	80.17	80.92	37.50
Classe 7	7.97	8.11	0.00
Classe 8	2.37	2.41	0.00
Classe 9	5.17	4.17	62.50
Classe 10	0.65	0.66	0.00






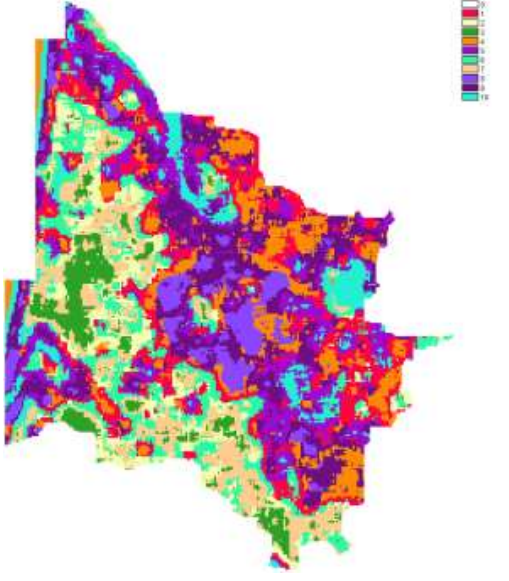
Dans l'image classée d'après les 5 indices construits avec CLC, trois classes contiennent la quasi-totalité de l'effectif de *Hyla meridionalis* : les classes 8 et 10 contiennent environ 75% de l'effectif. En ajoutant la classe 5, 92% de l'effectif est pris en compte. Dans celle classée d'après les 11 indices, les classes 4, 8 et 9 permettent de discriminer 86% de l'effectif.

Tab. LXXXI. Les classifications basées sur CLC distinguent la présence de *H. meridionalis* dans 3 types d'espace

Classification d'après les 3 composantes de l'ACP réalisée sur les 5 indices CLC (cf. Fig. 272)			
	Part des points dans les classes (%)		
	Toutes espèces	<i>Hyla meridionalis</i>	<i>Hyla molleri</i>
Classe 1	0.22	0.22	0.00
Classe 2	0.00	0.00	0.00
Classe 3	0.00	0.00	0.00
Classe 4	3.88	3.51	25.00
Classe 5 	17.67	17.98	0.00
Classe 6	0.00	0.00	0.00
Classe 7	0.00	0.00	0.00
Classe 8 	31.03	30.92	37.50
Classe 9	3.66	3.73	0.00
Classe 10 	43.53	43.64	37.50



Classification d'après les 3 composantes de l'ACP réalisée sur les 11 indices CLC (cf. Fig. 274 et Fig. 276)			
	Part des points dans les classes (%)		
	Toutes espèces	<i>Hyla meridionalis</i>	<i>Hyla molleri</i>
Classe 1	5.82	5.48	25.00
Classe 2	0.22	0.22	0.00
Classe 3	0.00	0.00	0.00
Classe 4 	35.34	35.96	0.00
Classe 5	2.37	2.41	0.00
Classe 6	4.31	3.73	37.50
Classe 7	0.65	0.66	0.00
Classe 8 	19.61	19.96	0.00
Classe 9 	31.47	31.36	37.50
Classe 10	0.22	0.22	0.00



Des traitements sont à poursuivre pour définir ces classes, autrement dit, caractériser le paysage des espaces contenant potentiellement des habitats pour les espèces de rainette. Enfin, l'intégration des différentes images dans un SDM permettra d'observer l'apport de cette nouvelle donnée environnementale au modèle global.

14.4 Discussion

Dans la perspective d'intégrer dans un modèle de répartition d'espèces (SDM) des informations sur l'habitat ou le paysage, l'utilisation d'indices paysagers calculés par le logiciel Fragstats sur la base de données d'occupation du sol préexistantes a été explorée. Les modes de fabrication de ces indices paysagers ont été démultipliés, selon la base de données d'occupation du sol utilisée et la taille de fenêtre coulissante. Dans un premier temps, il s'agit de questionner la compatibilité entre la résolution spatiale et la question de recherche : le lien entre l'espèce étudiée et son habitat. Cela amène dans un second temps à réfléchir aux enjeux du processus de modélisation.

Dans le logiciel Fragstats, la spatialisation des indices paysagers repose sur le principe de la fenêtre coulissante, qui attribue à la cellule centrale d'une fenêtre carrée, la valeur d'un indice calculé au niveau « paysage ». Ici, le paysage est défini par le périmètre de cette fenêtre et la cellule centrale est caractérisée en fonction de cet environnement (MCGARIGAL & MARKS, 1995). La taille de fenêtre correspond donc en un sens à une manière de borner le paysage en lui attribuant une superficie de manière arbitraire et liée à la taille de la cellule, autrement dit, de la résolution spatiale de la donnée d'occupation du sol.

Cette méthode permet donc de décrire la structure du paysage local. La démultiplication des tailles de fenêtres permet de discuter la taille du paysage et d'étudier dans quelle mesure celle-ci peut être un facteur explicatif, d'un point de vue écologique, de la présence des espèces, cette explicitation étant rarement réalisée (ANDRIEU, 2017). Il faut rappeler qu'au sens de l'écologie du paysage, « le paysage est défini comme une portion de territoire d'échelle kilométrique » (ANDRIEU, 2017). Cette définition peut questionner la nécessité de recourir à une résolution spatiale fine, supposée nécessaire vu le choix des espèces étudiées, qui disposent de faibles capacités de déplacement. Si l'hypothèse de départ était que l'utilisation d'indices paysagers, calculés à partir de données d'occupation du sol de résolution spatiale élevée (de l'ordre de la centaine de mètres) permettrait d'affiner la résolution des données climatiques de Météo-France (quelques kilomètres), il est intéressant de constater ce retournement de situation : le calcul d'indices paysagers avec le principe de la fenêtre coulissante permet finalement d'inclure des informations paysagères d'ordre kilométrique dans des cellules d'ordre hectométrique. Autrement dit, plutôt qu'inclure des données à haute résolution dans des données à plus basse résolution, ce sont des données de basse résolution qui ont été incorporées dans des données à haute résolution.

Par ailleurs, la comparaison entre les deux bases de données d'occupation du sol disponibles en Nouvelle-Aquitaine, CORINE Land Cover (CLC) et le Référentiel néo-aquitain d'occupation du sol à grande échelle (OCS-GE), vise à rendre visible et expliciter les différences de résultats liées à la résolution spatiale de fabrication de ces données. En effet, pour être utilisées dans Fragstats, ces bases de données, téléchargées au format vectoriel, ont été rastérisées selon une résolution spatiale qui correspond à la taille de l'unité minimale cartographiée : respectivement 100 m ($\sqrt{10}$ ha) et 500 m ($\sqrt{25}$ ha).

Or, l'objet observé varie en fonction de la résolution spatiale choisie (DRAGUT & EISANK, 2011). Ainsi, la complexité paysagère décrite par OCS-GE correspond au nombre plus important de polygones décrits dans cette base de données. Un parallèle peut être réalisé entre le fait que les indices de lisière apparaissent de bons prédicteurs de la présence de *Hyla molleri* lorsqu'ils sont calculés d'après le Référentiel néo-aquitain et de celle de *Hyla meridionalis* lorsqu'ils sont calculés d'après CORINE Land Cover. D'une certaine manière, la résolution spatiale aboutissant à des objets plus hétérogènes de

CORINE Land Cover semble correspondre aux caractéristiques écologiques de la rainette méridionale, plus généraliste que la rainette ibérique.

Le fait que des indices paysagers calculés dans de grandes fenêtres (5.5 km) et d'après la base de données de résolution spatiale la moins fine (CLC) soient identifiés en tant que prédicteurs de la présence des deux espèces de rainette, dont les capacités de déplacement sont faibles, peut paraître contre-intuitif. L'hétérogénéité qu'ils décrivent pourrait amener à supposer l'existence de plusieurs milieux juxtaposés et de petite taille. Cependant, de tels milieux sont mal décrits par les données utilisées. En effet, afin d'utiliser le logiciel Fragstats, les données d'occupation du sol, de format vectoriel, ont été transformées au format raster. Cette rasterisation a été effectuée par un paramétrage qui conserve, au sein de chaque cellule, le type d'occupation du sol dominant, c'est-à-dire celui qui occupe la plus grande superficie dans la cellule. Cette méthode a pour limite d'aboutir à une « disparition » des éléments linéaires, tels que les routes, qui peuvent pourtant avoir un impact non négligeable sur les rainettes, mais aussi des types d'occupation du sol dont la superficie est peu importante.

Avec un tel paramétrage, les discontinuités mises en exergue dans les indices paysagers calculés par Fragstats sont liées à la configuration des matrices Epaisseur de lisière et Contraste notamment et ne peuvent pas refléter la fragmentation des milieux par les routes, ou au contraire la proximité potentiellement bénéfique à de petits plans ou cours d'eau. Il serait donc intéressant de traduire en cartes les discontinuités modélisées d'après les paramètres du logiciel afin de mieux comprendre ce que sont capables d'exprimer les indices paysagers en termes de fragmentation et de connectivité des milieux. D'autre part, dans le précédent rapport, il était apparu que « les points de présence s'expliquent non pas par le poste dominant mais par les autres types d'occupation du sol présents dans ces mailles » (MALLARD, 2019). Tester la possibilité d'inclure la diversité d'une maille en une seule information pour la caractériser et l'inscrire dans le calcul des indices paysagers constitue l'une des perspectives de cette recherche méthodologique et appelle à se réinterroger sur les enjeux du processus de modélisation.

Plusieurs auteurs notent que le nombre de variables explicatives à implémenter dans le SDM est limité (Guisan & Zimmermann, 2000 ; Francesco Ficetola et al., 2014). La réduction du nombre de variables paysagères nécessite donc des choix, qui peuvent être réalisés, entre autres, selon des critères statistiques. Or, aucun choix réalisé *a priori* pour ce paramètre de taille de la fenêtre coulissante ne semble justifiable d'un point de vue statistique, ou plutôt, le choix de chacune des tailles de fenêtre est justifiable statistiquement (LOBRY, 2020).

Au même titre que les variables climatiques, les variables paysagères les plus pertinentes dépendent de l'espèce étudiée. Les variables paysagères sont diverses d'une part du point de vue de la caractéristique qu'elles quantifient (l'effet de lisière, la superficie, la fragmentation, ...) mais aussi diversifiées selon leur mode de construction (base de données d'occupation du sol utilisée, taille de fenêtre coulissante pour le calcul).

Pour le modèle de répartition (SDM), il s'agit de se demander si les variables à conserver permettent d'identifier la présence du groupe d'espèces pour tendre vers la modélisation d'un type de milieu – humide pour ces espèces – ou correspondent particulièrement à une seule espèce. Ainsi, les rasters en entrée du modèle peuvent varier.

Dans tous les cas, le modèle de répartition nécessitant un nombre réduit de variables, il apparaît nécessaire de combiner plusieurs variables paysagères en une variable synthétique. La modélisation a une puissance heuristique considérable, c'est un auxiliaire déterminant de la recherche même quand



Le modèle traduit la réalité de façon très imparfaite. Souvent, ce n'est pas le modèle qui est important mais la modélisation, le fait de construire le modèle. La méthode d'analyse d'images consistant à réaliser une classification à partir d'images de composantes principales permet de conserver une certaine généalogie de l'image et ainsi, l'explicitation du sens de la variable.

Afin que la modélisation joue son rôle d'aide à la décision politique, il paraît nécessaire de pouvoir en discuter les résultats mais aussi les modes de construction (CATTEAU & LOBRY, 2021a). Dans le contexte actuel de technicisation de l'environnement (GRANJOU, 2013 ; BOULEAU & DEUFFIC, 2016), la démultiplication des données environnementales au service de stratégies de suivi montre que leur construction même constitue un enjeu politique (CATTEAU & LOBRY, 2021b).

14.5 Conclusion

Les analyses présentées dans ce chapitre constituent la dernière étape vers la constitution d'un jeu de données paysagères intégrables en tant que variables explicatives dans les SDM. L'étude conserve une focale particulière sur les enjeux liés à la résolution spatiale des données et l'exploration des possibilités d'articuler des informations disponibles à différentes résolutions.

Dans un premier temps, il s'agissait d'analyser la capacité des métriques paysagères à prédire la présence chacune des espèces étudiées grâce à l'utilisation d'arbres de décision, afin de constituer un lot d'images d'indices pertinents vis-à-vis de l'espèce. Dans un second temps, les indices paysagers prédictifs ont été combinés avec une méthode d'analyse d'images qui permet de garder l'explicitabilité de la variable synthétique constituée.

Au fil du chapitre, il s'agit de démontrer tant les possibilités techniques de création d'indicateur paysager quantitatif que leurs enjeux théoriques quant à leur utilisation dans un modèle de répartition d'espèces. Si leur intégration en tant que variables explicatives dans un SDM paraît complexe, les indices paysagers visent à introduire, d'une part, le rôle joué par l'espace, et d'autre part, les interactions entre la société et la Nature et les dynamiques inhérentes au socio-écosystème qui peuvent être traduites en partie par l'occupation du sol.

L'intégration de la variable paysagère constituée dans un modèle de répartition d'espèce constitue la perspective de ce travail. Il s'agira d'observer si cette prise en compte des structures spatiales et des interrelations Nature-Société dans le modèle apparaît efficace et pertinente.