



HAL
open science

Améliorer les connaissances sur les processus écologiques régissant les dynamiques de populations d'auxiliaires de culture : modélisation couplant paysages et populations pour l'aide à l'échantillonnage biologique dans l'espace et le temps

Benoit Bellot

► **To cite this version:**

Benoit Bellot. Améliorer les connaissances sur les processus écologiques régissant les dynamiques de populations d'auxiliaires de culture : modélisation couplant paysages et populations pour l'aide à l'échantillonnage biologique dans l'espace et le temps. Sciences agricoles. Université de Rennes, 2018. Français. NNT : 2018REN1B008 . tel-01887527

HAL Id: tel-01887527

<https://theses.hal.science/tel-01887527>

Submitted on 4 Oct 2018

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.

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Bretagne Loire

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

*Mention : Statistiques/Modélisation en écologie, géosciences, agronomie et
alimentation*

Ecole doctorale EGAAL

Benoit Bellot

Préparée dans les unités de l'unité de recherche
U.M.R. INRA 1349 IGEPP et U.M.R. INRA 0980 BAGAP

**Améliorer les
connaissances sur les
processus écologiques
régissant les dynamiques
de populations
d'auxiliaires de culture :
Modélisation couplant
paysages et populations
pour l'aide à
l'échantillonnage dans
l'espace et le temps**

Thèse dirigée par :

Jacques BAUDRY

Directeur de Recherche, INRA Rennes / *Directeur*

Nicolas PARISEY

Ingénieur de Recherche, INRA Rennes / *Co-directeur*

Sylvain POGGI

Chargé de Recherche, INRA Rennes / *Co-directeur*

Thèse rapportée par :

Florence LE BER

Directrice de Recherche, ENGEES Strasbourg

Aude VIALATTE

Maître de Conférence, ENSAT Toulouse

et soutenue à Rennes le 18 Avril 2018

devant le jury composé de :

Florence LE BER

Directrice de Recherche, ENGEES Strasbourg / *Rapporteuse*

Thomas DELATTRE

Chargé de recherche, INRA Avignon / *Examineur*

Anne LE RALEC

Professeure, Agrocampus Ouest Rennes / *Présidente du jury*

Jean-Sébastien PIERRE

Professeur émérite, CNRS Rennes / *Examineur*

Benoit RICCI

Chargé de recherche, INRA Dijon / *Examineur*

Jacques BAUDRY

Directeur de Recherche, INRA Rennes / *Directeur de thèse*

Améliorer les connaissances sur les
processus écologiques régissant les
dynamiques de populations d'auxiliaires de
culture

*Modélisation couplant paysages et populations pour l'aide à
l'échantillonnage biologique dans l'espace et le temps*

Auteur : Benoit Bellot

Encadré par :

Jacques Baudry, Ph.D.

N. Parisey, Ph.D.

S. Poggi, Ph.D.

Réalisée à l'Université de Rennes 1
Rennes, France

13 février 2018

Résumé

Une alternative prometteuse à la lutte chimique pour la régulation des ravageurs de culture consiste à favoriser les populations de leurs prédateurs en jouant sur la structure du paysage agricole. L'identification préalable de structures spatio-temporelles favorables aux ennemis naturels peut se faire par l'exploration de scénarios paysagers via une modélisation couplée de paysages et de dynamiques de population. Dans cette approche, les dynamiques de populations sont simulées sur des paysages virtuels aux propriétés structurales contrôlées, et l'observation des motifs de populations associés permet l'identification de structures favorables. La modélisation des dynamiques de populations repose cependant sur une connaissance fine des processus écologiques (*e.g.* dispersion ou démographie) et de leur variabilité entre les différentes unités du paysage. L'état actuel des connaissances sur les mécanismes écologiques régissant les dynamiques des ennemis naturels de la famille des carabidés demeure l'obstacle majeur à la recherche *in silico* de scénarios paysagers favorables.

La littérature sur les liens entre motifs de population de carabes et variables paysagères permet de formuler un ensemble d'hypothèses en compétition sur ces mécanismes. Réduire le nombre de ces hypothèses en analysant les convergences entre les motifs de population qui leur sont associés, et étudier la stabilité de ces convergences le long d'un gradient paysager apparaît comme une première étape nécessaire vers l'amélioration de la connaissance sur les processus écologiques. Dans une première partie, nous proposons une heuristique méthodologique basée sur la simulation de modèles de réaction-diffusion porteurs de ces hypothèses en compétition. L'étude des motifs de population a permis d'effectuer une typologie des modèles en fonction de leur réponse à une variable paysagère, via un algorithme de classification basé sur les *K-médoïdes*. Cette classification a permis de déterminer un nombre restreint de modèles sur la base de leurs réponses à la variable paysagère.

La sélection de l'hypothèse écologique la plus plausible parmi cet ensemble irréductible doit s'effectuer sur la base d'une observation des motifs de population sur le terrain. Cela implique que ces derniers soient caractérisés à des résolutions spatiales et temporelles suffisantes pour sélectionner une unique hypothèse parmi celles en compétition. Dans la deuxième partie, nous proposons une heuristique méthodologique permettant de déterminer *a priori* des stratégies d'échantillonnage maximisant la robustesse de la sélection d'hypothèses écologiques. Dans un premier temps, la simulation de modèles de réaction-diffusion représentatifs des hypothèses écologiques en compétition permet de générer des données biologiques virtuelles en tout point de l'espace et du temps. Ces données biologiques sont ensuite échantillonnées suivant des protocoles différant dans l'effort total d'échantillonnage, le nombre de dates, le nombre de points par unité d'espace et le nombre de réplicats de paysages. Les motifs des populations sont caractérisés à partir de ces échantillons. Le potentiel des stratégies d'échantillonnage est évalué via un algorithme de classification qui classe les modèles biologiques selon les motifs de population associés. L'analyse des performances

de classification, *i.e.* la capacité de l'algorithme à discriminer les processus écologiques, permet de sélectionner un protocole d'échantillonnage optimal. Nous montrons également que la manière de distribuer l'effort d'échantillonnage entre ses composantes spatiales et temporelles est un levier majeur sur l'inférence des processus écologiques.

La réduction du nombre d'hypothèses en compétition et l'aide à l'échantillonnage pour la sélection de modèles répondent à un besoin fort dans le processus d'acquisition de connaissances écologiques pour l'exploration *in silico* de scénarios paysagers favorisant des services écosystémiques. Nous discutons dans une dernière partie des implications de nos travaux et de leurs perspectives d'amélioration.

Mots clés : *Écologie du paysage ; Services écosystémiques ; Régulation de ravageurs ; Modélisation de paysages ; Dynamiques de population ; Réaction-diffusion ; Sélection de modèles ; Inférence des processus écologiques ; Algorithme de classification ; Échantillonnage spatio-temporel ; Carabidés*

Table des matières

Résumé (<i>abstract</i>)	I
Table des matières	IV
Liste des figures	VI
I Introduction générale	1
1.1 Paysages, processus, motifs	2
1.2 Les limites des approches corrélatives pour relier processus et motifs .	7
1.3 Inférence des processus écologiques à partir des motifs de population .	8
1.4 Positionnement et contenu de la thèse	10
1.5 Considérations personnelles sur la démarche employée	13
Références	17
II Étude de l'effet de l'hétérogénéité de configuration du paysage sur les dynamiques de population dans un contexte de connaissance limitée sur les processus écologiques	19
1 Présentation de la démarche, du modèle biologique et de la variable paysagère	20
1.1 Démarche méthodologique	20
1.2 Présentation du modèle biologique	23
1.3 L'hétérogénéité de configuration du paysage : définition, pertinence et métriques	25
2 Matériel et Méthode	28
2.1 Génération de paysages agricoles virtuels	28
2.2 Typologie des processus en fonction des motifs	33
3 Résultats et discussion	40
3.1 Génération des paysages virtuels	40
3.2 Typologie des modèles de processus sur la base de leur signature populationnelle	41
3.3 Relation entre les métriques de population et la variable paysagère . .	43
4 Conclusion du chapitre 2	47
Références	52

III	Inférence des processus écologiques par l'analyse des motifs de population : un guide méthodologique pour l'échantillonnage dans l'espace et le temps	53
1	Préface à l'article	54
1.1	Présentation générale de la démarche	54
1.2	Remarque sur les hypothèses écologiques sélectionnées	56
2	Article soumis à <i>Ecological Modelling</i>	57
2.1	Introduction	59
2.2	Material and Methods	61
2.3	Results	70
2.4	Discussion and conclusions	81
3	Perspectives d'amélioration concernant le mode d'échantillonnage	85
	Références	94
IV	Discussion générale et conclusion	95
4.1	Bilan des travaux	96
4.2	Lien entre processus écologiques et services écosytémiques	101
4.3	Conclusion	104
	Références citées dans la thèse	116
	Appendices	117
A	Annexes relatives au chapitre 2	118
B	Annexes relatives à l'article	120
B.1	Process models calibration	120
B.2	Numerical scheme	121
B.3	Population metrics computed after spatio-temporal sampling	122
B.4	Global sensitivity analysis	124

Table des figures

1.1	Phénomènes influant la distribution des individus à l'échelle du paysage	4
1.2	Lien entre hétérogénéité structurale, processus écologiques et motifs de population	5
1.3	Table 1 : Définition de termes communément utilisés en écologie du paysage	6
1.4	Positionnement des travaux	12
1.1	Présentation de la démarche méthodologique (chapitre 2)	23
1.2	Hétérogénéité de configuration et hétérogénéité de composition	26
2.1	Exemples de tessellations	29
2.2	Zone Atelier Armorique	32
2.3	Table 2 : Formulation des modèles de Réaction-diffusion	36
3.1	Exemples de paysages générés	41
3.2	Écart entre valeurs cibles et valeurs atteintes pour les paysages générés	41
3.3	Regroupement en écotypes et variables de population associées	44
3.4	valeurs des métriques de population des modèles médoïdes le long du gradient de la variable paysagère	45
3.5	Population moyenne au cours du temps	46
2.1	Scheme of the heuristic framework	63
2.2	table 3 : Process Models formulation	73
2.3	table 4 : Parameter reference values and ranges for the Global Sensitivity Analysis	74
2.4	Classification performances according to the total sampling effort	75
2.5	Contour plots of predicted classification accuracies by the General Additive Model	75
2.6	Pairwise confusion rates	76
2.7	Confusion matrix associated with the selected sampling design	77
2.8	Sensitivity analysis results	79
2.9	Pairwise type I classification errors	80
3.1	Distribution spatiale de la métrique de contraste	86
A.1	Représentation des clusters pour chaque valeur d'hétérogénéité du paysage	119
B.1	Table 5 : Population metrics computed after sampling	123

B.2 Table 6 : Population metrics computed after sampling (continued) . . 124

Première partie

Introduction générale

1.1	Paysages, processus, motifs	2
1.2	Les limites des approches corrélatives pour relier processus et motifs .	7
1.3	Inférence des processus écologiques à partir des motifs de population .	8
1.4	Positionnement et contenu de la thèse	10
1.5	Considérations personnelles sur la démarche employée	13
	Références	17

1.1 Paysages, processus, motifs

Selon que l'observateur est un éleveur, un artiste peintre ou un enfant, ce qui va être perçu d'un même espace va être radicalement différent. L'un pourra en avoir une vision fonctionnelle, percevant les points d'eaux et les zones de pâturage, l'autre en observera les contrastes de couleur, de lumière et l'autre encore se demandera où il pourra se cacher ou fabriquer une cabane. Leur point commun est qu'ils percevront tous trois l'hétérogénéité de cet espace et qu'ils apprécieront, consciemment ou non et selon leurs filtres cognitifs et émotionnels, la qualité de chaque élément le constituant.

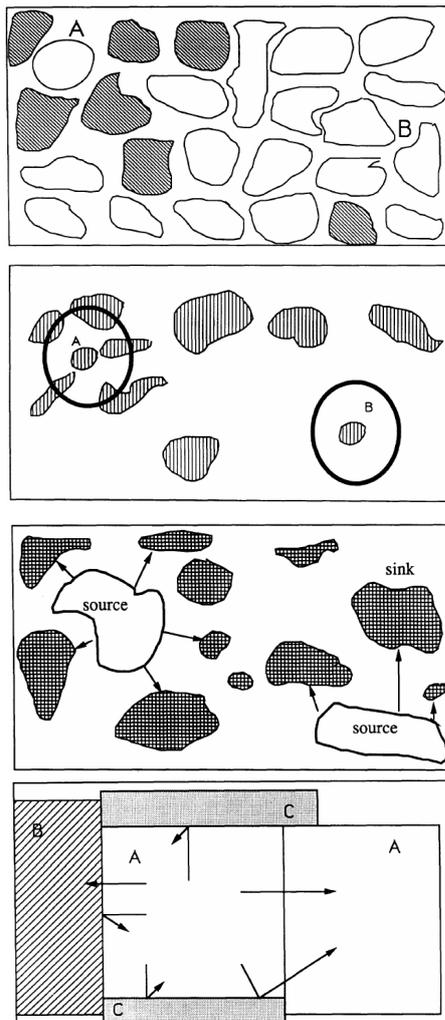
Malgré la diversité des définitions qui lui sont attribuées, et quelle que soit la discipline concernée, cette notion d'hétérogénéité est indissociable de la notion de paysage. Dans le domaine de l'écologie du paysage, partant de notions plutôt abstraites pour se définir de plus en plus concisément, le paysage a tantôt été, selon les auteurs, « *une portion de région hétérogène composée d'une association particulière d'écosystèmes* (Forman et Godron 1986) », « *une mosaïque constituée d'éléments divers répétés* » (Forman 1995b), ou encore « *un niveau d'organisation des systèmes écologiques, supérieur à l'écosystème ; caractérisé essentiellement par son hétérogénéité et par sa dynamique gouvernée pour partie par les activités humaines* » (Burel et Baudry 1999).

Une définition plus concise formulée par (Forman 1995a) et qu'on retrouve couramment jusqu'à aujourd'hui (Turner et Gardner 2015), désigne le paysage comme « *un espace hétérogène en au moins une variable d'intérêt* ». Nous complétons cette définition en précisant que la variable d'intérêt peut-être quantitative et/ou qualitative, et qu'elle peut être mesurée d'une manière objective (*e.g.* type de couvert, caractéristiques édaphiques, conditions (micro-)climatiques). Le paysage peut donc être considéré comme une mosaïque hétérogène dont chacun des éléments constitutifs est décrit par une ou plusieurs variables, ou attributs (Duelli 1997; Fahrig, Baudry *et al.* 2011; Vasseur *et al.* 2013). Nous qualifierons de *patch*¹ chaque unité du paysage présentant une homogénéité en chacun de ses attributs. Le paysage possède une *échelle* spatiale, qui est la combinaison de l'*étendue* (la taille du paysage) et de sa *résolution* (la taille de son plus petit élément constitutif). Défini ainsi, le paysage est un espace structurel et objectif. Sa dimension fonctionnelle n'apparaît que lorsque l'on décrit une relation entre ses éléments constitutifs et les processus écologiques. Fahrig, Baudry *et al.* (2011) font ainsi la distinction entre l'hétérogénéité "structurelle", où les éléments du paysage sont identifiés par leurs caractéristiques physiques, sans référence à une espèce ou un groupe d'espèces, et l'hétérogénéité "fonctionnelle" où ils sont caractérisés par la teneur et la qualité des ressources

1. Les termes en italiques présentés dans cette l'introduction sont résumés dans la table 1.3

qu'ils apportent à une espèce/groupe d'espèces. La notion d'hétérogénéité fonctionnelle permet de mieux appréhender la notion du paysage selon [Wiens \(1999\)](#), qui le définit comme « *l'échelle spatiale à laquelle des patterns spatiaux influencent des processus écologiques* ».

Ce passage d'un support inerte à un support fonctionnel fait intervenir la notion de processus écologique. [Schröder et Seppelt \(2006\)](#) définissent la notion de processus écologique comme l'interaction de différents objets dans le paysage. Cela inclut à la fois les interactions entre les entités biologiques et leur environnement (*e.g.* consommation de ressources, mouvement), mais également les interactions entre les entités biologiques (compétition intra/interspécifique, prédation). Ici le paysage devient une notion espèce-, voire individu-spécifique. Chaque élément du paysage sera perçu par un organisme selon plusieurs facteurs, comme la quantité et la qualité des ressources disponibles, la facilité de déplacement ou les risques de prédation en son sein. En d'autres termes, à chaque élément du paysage sont associés des facteurs jouant sur la démographie et le mouvement de l'organisme d'intérêt. La quantité et l'agencement des éléments du paysage conditionnent la distribution de la population en son sein via quatre phénomènes, détaillés en figure 1.1 : la complémentarité, la supplémentation, les dynamiques sources-puits, les effets de voisinage.



Complémentation : une espèce a besoin de ressources non substituables trouvées dans les deux types de couverts (patches blancs et patches ouverts). Les zones du paysage où les deux types de couverts sont relativement proches supporteront plus d'individus que des zones où l'un des habitats est plus rare.

Supplémentation : Une espèce a besoin de ressources trouvées dans les habitats grisés. Les patches A et B sont trop petit pour supporter une population seuls. Cependant les patch A peut supporter une population si les individus sont capables de prélever les ressources dans les patches environnants

Dynamiques sources-puits : une espèce occupe deux types d'habitats : type de patch de bonne qualité capable de produire des individus en excès, et un type de patch de moindre qualité qui ne produit pas assez d'individus pour maintenir une population locale. Les populations dans les puits sont dépendantes des individus immigrants des patches sources voisins.

Effets de voisinage : La capacité des disperseurs à coloniser les patches voisins est conditionnée par la nature des interfaces. L'habitat C ne peut être colonisé par les individus de l'habitat A du fait de la nature imperméable de l'interface entre les deux types de couverts

FIGURE 1.1 – Les quatre phénomènes influant la distribution des individus à l'échelle du paysage. Adapté de [Dunning et al. \(1992\)](#).

In fine, l'hétérogénéité structurale du paysage, transformée par le filtre des processus en hétérogénéité fonctionnelle, est responsable de la distribution de la population dans l'espace et le temps. Autrement dit, l'hétérogénéité spatiale des processus due à l'hétérogénéité structurale du paysage, génère des motifs spatio-temporels de population au sein du paysage (Fig. 1.2). On définit ainsi un *motif* (*pattern* en anglais, terme également employé dans la suite du document) comme toute observation de l'expression de l'interaction entre hétérogénéité du paysage et processus écologiques ([Vinatier 2010](#)). Selon cette définition, le pattern peut être décrit à différentes résolutions, allant de mesures fines d'autocorrélations spatiales et/ou temporelles en tout point de l'espace et du temps ([Fortin et Dale 2005](#) ; [Schröder et Seppelt 2006](#)) à des descripteurs globaux de la population (*e.g.* moyenne de population agrégée dans le temps à l'échelle du paysage). Au cours de ce manuscrit, nous utiliserons également le terme de « signature », porteur du lien spécifique entre le processus et

son observation, pour parler de *patterns*.

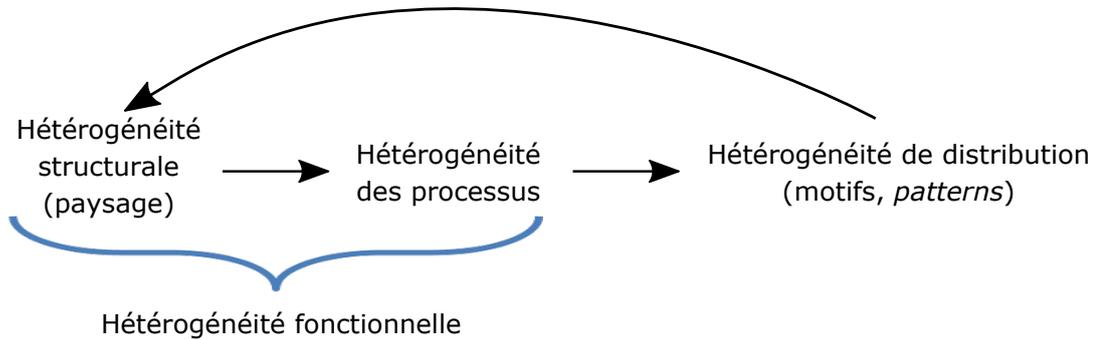


FIGURE 1.2 – Relation entre hétérogénéité structurale (composition du paysage et agencement dans l’espace et le temps de ces éléments), processus (valeurs des traits de vie au sein des éléments du paysage) et motifs de population.

La flèche qui retourne des *patterns* à la structure du paysage dans la figure 1.2 signifie que ceux-ci peuvent avoir un effet rétroactif sur la structure paysagère. Par exemple une forte densité d’individus, initialement due à une forte quantité de ressources en une zone du paysage (caractéristique structurale), rétroagira sur le paysage par épuisement de la ressource, avec pour effet la modification de la distribution des individus.

C’est ce triptyque paysage-processus-pattern qui est au cœur de l’écologie du paysage. En modifiant la structure du paysage (la quantité de ressources et la manière dont elles sont agencées spatialement et temporellement), il est possible de jouer par effet de levier sur les motifs de l’objet biologique dans l’espace et le temps (distribution des individus, des populations, structure de la communauté). Ce levier est utilisé à de nombreuses fins en écologie du paysage : aménagement de réserves pour la conservation et la gestion de la biodiversité (Moloney et Levin 1996 ; Aben *et al.* 2016), aménagement du paysage pour le contrôle d’espèces invasives (Jongejans *et al.* 2008 ; Savage et Renton 2014), recherche de paysages maximisant des services écosystémiques (Landis *et al.* 2000 ; Jonsson *et al.* 2014 ; Parisey *et al.* 2016 ; Bourhis, Poggi, Mammeri, Le Cointe *et al.* 2017).

FIGURE 1.3 – Définition de termes communément utilisés en écologie du paysage (adapté de Turner et Gardner (2015)).

Terme	Définition
Composition	Quels types d’habitats ou de couverts sont présents, et en quelles quantités
Configuration	Arrangement des éléments spatiaux ; souvent utilisée comme synonyme de la structure spatiale ou de la structure des patches
Type de couvert (Cover type)	Catégorie qui distingue les différents habitats, écosystèmes, ou types de végétation dans un paysage
Interface (Edge)	Frontière entre deux types de couverts différents
Fragmentation	Morcellement d’un type d’habitat ou d’un type de couvert en parcelles plus petites et déconnectées les unes des autres ; souvent associée, mais non équivalente, à la perte d’habitat
Hétérogénéité (Heterogeneity)	La qualité ou l’état d’être constitué de différents éléments, comme lorsque des habitats ou des types de couverts différents sont présents dans un paysage ; opposé de l’homogénéité, dans laquelle les éléments sont identiques
Motif (pattern)	Observation d’un phénomène, susceptible de contenir des informations concernant les mécanismes desquels elle émerge
Paysage (Landscape)	Une surface qui est spatialement hétérogène pour au moins un facteur d’intérêt, qui peut être mesuré d’une manière objective
Processus	Interaction de différents objets dans le paysage. Nous l’utiliserons pour décrire plus spécifiquement les valeurs des traits de vie au sein des différents éléments du paysage
Patch	Une surface homogène qui diffère de son environnement voisin en nature ou en apparence
Échelle (scale)	Dimension spatiale ou temporelle d’un objet ou d’un processus, caractérisée à la fois par le grain (la taille d’une unité d’observation) et l’étendue

1.2 Les limites des approches corrélatives pour relier processus et motifs

Comme nous l'avons vu précédemment, l'objectif appliqué de l'écologie du paysage est d'identifier des structures paysagères qui ont une influence, positive ou négative, sur un organisme d'intérêt. Une approche assez intuitive, et qui a largement été employée depuis l'avènement de l'écologie du paysage, est donc de comparer directement les relations entre une variable paysagère d'intérêt avec des descripteurs de l'espèce ou du peuplement (*e.g.* Aviron *et al.* 2005 ; Chaplin-Kramer *et al.* 2011 ; Eyre *et al.* 2013 ; Puech *et al.* 2014). Dans ce type d'approche, que nous qualifierons de corrélative, descriptive, ou encore phénoménologique, les relations entre les variables paysagères et les descripteurs biologiques sont analysées à l'aide d'outils statistiques. L'espèce d'intérêt est échantillonnée dans une diversité de paysages, présentant un gradient de la variable paysagère d'intérêt. Une relation, significative ou non, linéaire ou non, est alors établie, et les conclusions apportées à ce type d'étude sont par exemple que « *la taille moyenne des parcelles a un effet négatif sur les mesures de biodiversité* » (Fahrig, Girard *et al.* 2015), ou encore que « *les coccinelles sont plus abondantes dans les paysages présentant une forte proportion de forêts et de prairies que dans les paysages dominés par des surfaces agricoles* » (Gardiner *et al.* 2009). Ce type de conclusions est bénéfique à au moins deux égards : *i*) d'une part, elles apportent des réponses accessibles et facilement transmissibles aux interlocuteurs « non scientifiques » qui peuvent avoir un pouvoir de décision sur l'aménagement du paysage (*e.g.* coopératives agricoles) et *ii*) d'autre part, ce type de conclusions sont souvent accompagnées d'hypothèses sur les processus écologiques responsables des observations (*e.g.* « *les résultats suggèrent que les mouvements des individus entre les parcelles expliquent mieux les distributions des populations que des différences dans le taux de survie* » (Marrec *et al.* 2015)).

Si les études corrélatives en écologie du paysage sont nécessaires car elles permettent de bâtir de nouvelles hypothèses ou d'apporter du crédit à des hypothèses existantes sur les processus écologiques, en revanche leurs conclusions sont à prendre avec précaution avant d'être utilisées dans un cadre d'aménagement du paysage. D'abord, comme le rappellent Turner et Gardner (2015), la corrélation entre des structures spatiales et des motifs de population n'implique pas une relation de causalité. Par exemple, Fahrig, Girard *et al.* (2015), dans leur étude mettant en évidence une corrélation négative entre la taille des parcelles et la biodiversité, soulignent que l'application de produits phytosanitaires n'a pas pu être mesurée. La corrélation négative entre taille des parcelles et biodiversité pourrait provenir d'une moindre application de produits phytosanitaires dans les paysages à petite parcelles. Préco-

niser des aménagements du paysage à partir de corrélations directes entre structure et motifs de population peut ne pas avoir l'effet escompté si des variables non mesurées ont un meilleur pouvoir explicatif que celles qui l'ont été (Turner et Gardner 2015). En somme, un modèle basé sur des corrélations peut avoir un très bon pouvoir prédictif sur le système particulier dont il est issu, mais ne pas être adapté à un autre système, i.e. avoir un faible pouvoir de généralisation (McIntire et Fajardo 2009 ; Jongejans *et al.* 2008). McIntire et Fajardo (2009) comparent l'approche corrélatrice au fait de commenter une course de chevaux *a posteriori*, où il est facile d'expliquer pourquoi tel cheval a gagné (*belle de nuit*, la jument la plus rapide, est tombée pendant la course), mais n'est pas généralisable à une autre course (*belle de nuit* sera de retour sur le circuit).

La généralisation du lien entre structure du paysage et motifs de population doit passer par une prise en compte explicite des processus (Jongejans *et al.* 2008). Formellement, les mécanismes écologiques (*e.g.* la vitesse et la direction du déplacement ou le taux de mortalité dû à la prédation) doivent être définis en chacun des éléments constituant le paysage. Les motifs de populations sont alors une caractéristique émergente de l'interaction entre la quantité et la configuration des éléments du paysage et les valeurs des mécanismes au sein de chacun de ces éléments paysagers (Peck 2004 ; Grimm *et al.* 2005). La robustesse des prédictions de l'effet de la structure paysagère repose donc sur une connaissance des mécanismes écologiques et leur hétérogénéité au sein des unités du paysage.

Malheureusement, l'acquisition de cette connaissance fine des processus en lien avec leur environnement est une tâche complexe, à tel point que l'étude du lien entre processus et motifs peut définir en soi la discipline de l'écologie (Tilman et Kareiva 1997 ; Tuda 2007). Les échelles temporelles et spatiales inhérentes à l'objet paysage, son caractère unique et la complexité des processus écologiques qu'il héberge rendent caduque l'expérimentation sur le terrain comme seule méthode pour déterminer la nature causale des processus sur les motifs de population (Ricci *et al.* 2013 ; Turner et Gardner 2015). La combinaison de méthodes impliquant des allers-retours entre terrain et modélisation permettront de mieux comprendre la relation processus-pattern et à terme, de prédire l'effet de la structure du paysage sur les motifs de population.

1.3 Inférence des processus écologiques à partir des motifs de population

La plupart du temps, lorsque l'on cherche à étudier l'effet de la structure du paysage sur un organisme, on dispose déjà d'hypothèses sur les processus écolo-

giques, issues d'interprétations de motifs de population à l'échelle du paysage (*e.g.* Marrec *et al.* 2015 ; Fahrig, Girard *et al.* 2015), d'expérimentations en conditions semi-contrôlées (Allema, Werf, Lenteren *et al.* 2014), de discussions avec des experts ou même d'intuitions personnelles. Ces hypothèses sont pour la plupart partielles, incertaines et parfois en contradiction. Au lieu d'être en possession d'une unique hypothèse solidement étayée, on dispose plutôt d'un ensemble d'hypothèses incertaines que l'on cherche à infirmer ou à confirmer (Holling et Allen 2002). Cette diversité des possibles concernant la relation entre l'organisme et son environnement présage d'une diversité des possibles dans l'effet de la structure du paysage sur les dynamiques de population qui ne peut être satisfaisante lorsque l'on cherche à utiliser le paysage comme un levier pour la gestion d'espèce.

Cependant avant de se lancer dans la complexe question de l'inférence des processus écologiques, il est légitime de s'interroger sur l'utilité de l'acquisition de ces connaissances en vue des objectifs que l'on poursuit en termes de gestion de population de l'organisme : dans quelle mesure des hypothèses en compétition, porteuses de processus écologiques différents, amènent elles à des motifs différents ? S'il est possible de mettre en lumière que les motifs de population issus de processus différents sont similaires quelle que soit la valeur de la variable paysagère d'intérêt, alors il n'est pas nécessaire d'améliorer les connaissances sur les processus pour statuer sur l'effet de la variable sur l'organisme (Peck 2004).

S'il s'avère en revanche que les processus différents amènent à des motifs différents, alors la connaissance doit être améliorée : il est nécessaire de savoir quels processus sous-tendent les dynamiques de population de l'organisme d'intérêt pour pouvoir, à terme, être capable de prédire les effets d'aménagements paysagers sur cet organisme.

Depuis la fin des années 1990, un consensus méthodologique s'établit progressivement sur la démarche à suivre pour inférer de manière robuste les processus écologiques à partir des motifs de population (Grimm *et al.* 2005 ; Holling et Allen 2002 ; McIntire et Fajardo 2009 ; Vinatier *et al.* 2011). La démarche s'articule autour de trois piliers :

1. La formulation a priori d'hypothèses multiples sur les processus écologiques
2. La traduction de ces hypothèses en modèles mécanistes
3. La comparaison des motifs de population issus de la simulation spatialement explicite des modèles et des motifs observés sur le terrain

Les hypothèses qui ne donnent pas des observations similaires à l'existant peuvent être rejetées. Si les hypothèses qui n'ont pas été rejetées sont considérées comme étant en adéquation suffisantes avec les observations, on peut considérer qu'on a capturé les mécanismes cruciaux expliquant les motifs de population, et qu'il n'est

alors pas nécessaire d’approfondir les connaissances sur les mécanismes sous-jacents (Peck 2004 ; Grimm *et al.* 2005 ; Winsberg 2003). Cette démarche occupe une place centrale dans la discipline émergente qu’est l’écologie du mouvement (Nathan 2008) et a permis d’identifier, par exemple, les mécanismes de mouvement et d’utilisation de la ressource responsables des motifs de vol de papillons (McIntire, Rompré *et al.* 2013).

Le troisième point, *i.e.* la comparaison des motifs simulés et observés, soulève la problématique de l’acquisition de données biologiques sur le terrain. La question de l’échantillonnage est centrale, car une mauvaise conception du protocole d’acquisition des données peut aboutir à une description trop grossière des motifs de population pour permettre la discrimination entre les hypothèses biologiques (Zurell *et al.* 2010). Cette question de recherche de stratégies d’échantillonnage dans une optique d’inférence des processus écologiques, bien que cruciale, n’a été que peu abordée (Schweiger *et al.* 2016). Dans le domaine de l’écologie du paysage, il n’existe à notre connaissance aucune heuristique qui permet de déterminer *a priori* des stratégies d’échantillonnage maximisant la robustesse de l’inférence des processus écologiques à partir des motifs de population. Nous nous proposons de pallier cette carence méthodologique en construisant une heuristique pour la sélection de stratégies d’échantillonnage.

1.4 Positionnement et contenu de la thèse

Ce travail de thèse s’inscrit dans la démarche générale de l’acquisition de connaissance sur le lien entre processus écologiques et motifs de populations. Nous abordons, en accord avec la démarche "d’inférence inductive", la notion de "connaissance limitée" par la multiplicité des hypothèses sur les mécanismes qui sous-tendent les dynamiques de population de l’organisme d’intérêt. Le manuscrit s’articule autour des deux problématiques suivantes :

1. Une connaissance limitée sur les processus écologiques est-elle un obstacle à l’établissement d’un lien entre la structure du paysage et les motifs de la population d’intérêt ? En d’autres termes, les processus écologiques en compétition génèrent-ils des motifs de population suffisamment similaires pour ne pas avoir besoin de trancher entre les hypothèses écologiques ?
2. Si les hypothèses sur les processus amènent à des motifs de population différents, comment est-il possible de sélectionner la (les) hypothèse(s) la(les) plus plausible(s) sur la base des motifs de population, en prenant en compte que la description des motifs est tributaire de la stratégie d’échantillonnage sur le terrain ? En d’autres termes, quelle stratégie d’échantillonnage est la plus à

même de nous permettre de trancher entre les différentes hypothèses sur les processus sous-jacents ?

Chacun des deux chapitres centraux de la thèse présente et illustre un cadre méthodologique pour aborder ces problématiques. Les hypothèses écologiques sont traduites en modèles de dynamiques de populations spatialement explicites, où chaque modèle représente une des alternatives possibles du « comportement » de l'organisme d'intérêt au sein du paysage. Ces modèles alternatifs seront dénommés « modèles de processus » (abrévés *PM* pour *Process Models*) tout au long du manuscrit. La figure 1.4 présente la manière dont ce travail de thèse s'inscrit dans le processus plus large d'acquisition de connaissances sur le lien processus-motifs et son application en écologie du paysage.

Dans le chapitre 2, nous présentons et illustrons un cadre méthodologique permettant d'évaluer l'effet d'une variable paysagère sur les dynamiques de population d'un organisme d'intérêt, en prenant en compte la diversité des processus écologiques possibles. Une typologie des modèles selon leur réponse à la variable paysagère est effectuée. L'ensemble initial d'hypothèses sur les processus écologiques est alors réduit à un ensemble de modèles mécanistes qui génèrent des réponses différentes à la variable paysagère.

Le chapitre 3 propose et illustre une heuristique méthodologique pour l'échantillonnage des données biologiques dans l'espace et dans le temps dans un objectif de sélection des processus écologiques parmi cet ensemble irréductible d'hypothèses. Enfin, le dernier chapitre apporte un regard critique sur la démarche mise en place durant ce travail de thèse en abordant ses limites, ses implications et les perspectives que nous en concevons à court et à moyen termes.

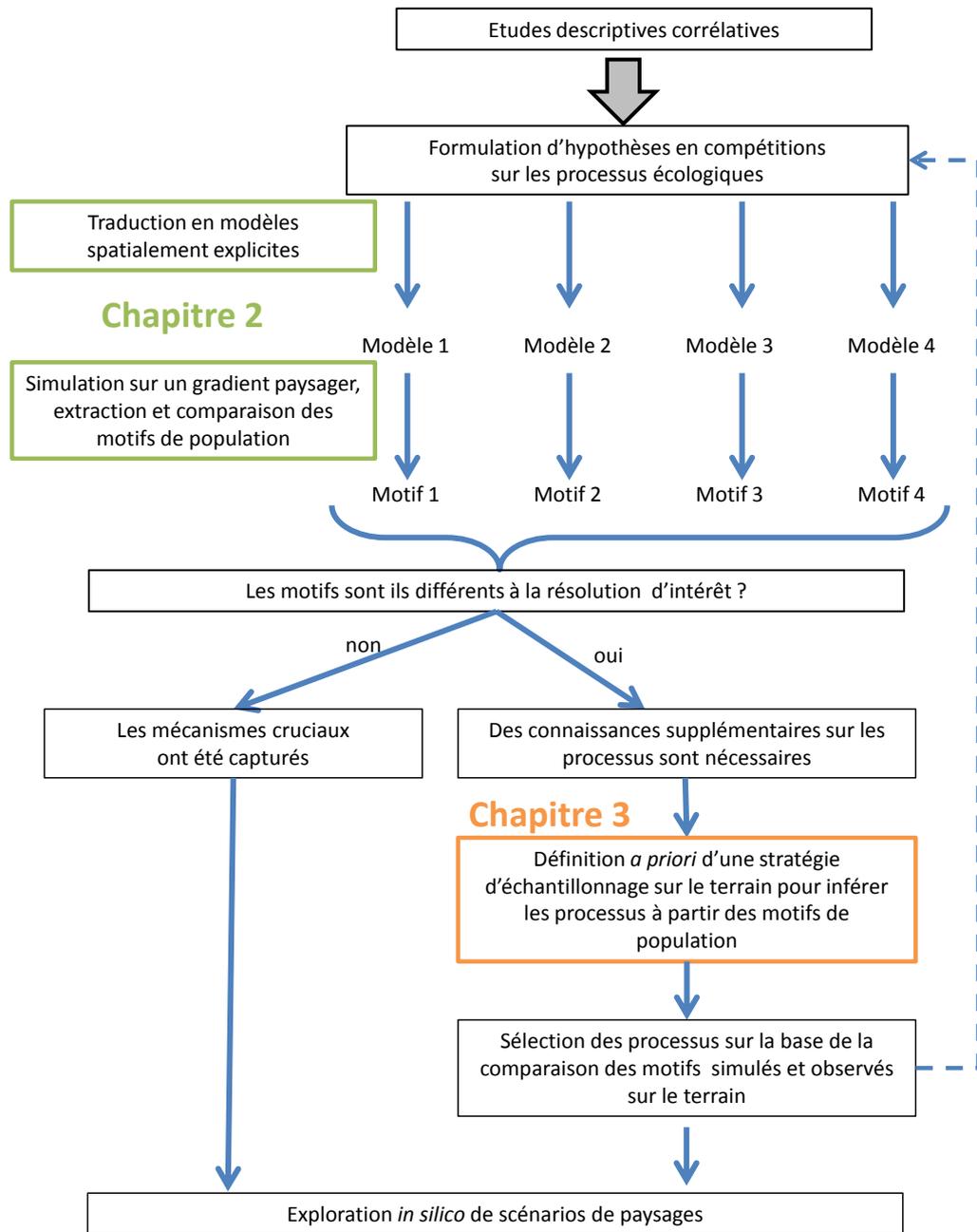


FIGURE 1.4 – Positionnement des travaux de thèse dans la démarche de caractérisation du lien entre processus écologiques et motifs de population, dont l’objectif appliqué est l’identification, via l’exploration *in silico* de scénarios d’aménagements, de structures favorables à la gestion d’une espèce ou d’une communauté.

1.5 Considérations personnelles sur la démarche employée

Du fait de l'impossibilité de traiter rigoureusement dans le temps imputable à un travail de thèse l'ensemble des étapes en amont (acquisition de connaissance pour l'établissement des hypothèses, calibration des modèles sur la base de données) et en aval (sélection de la bonne hypothèse sur les processus par un échantillonnage sur le terrain), nous nous sommes concentrés sur la partie centrale qu'est la modélisation mécaniste des processus et leur lien avec les motifs de population. Cela fait de cette thèse un travail résolument méthodologique, présentant une manière de traiter le manque de connaissance sur les processus écologiques plutôt que l'apport de connaissances arrêtées sur le lien process-pattern d'un organisme en particulier. Cette démarche a été motivée par la prise de conscience progressive de l'impossibilité de traiter rigoureusement dans le temps imparti un travail de modélisation et un travail de terrain pour une acquisition ciblée de données permettant de valider le modèle.

Cette prise de conscience, quoique tardive et douloureuse, de l'impossibilité de fournir un modèle mécaniste (encore un ?) modélisant les dynamiques de population d'un organisme, avec des paramètres calibrés sur un jeu de données biologiques unique (que nous aurions certainement jugé *a posteriori* insuffisant pour assurer la robustesse et la capacité de généralisation de nos conclusions) nous a permis de prendre du recul et de poser ces nouvelles problématiques.

Références de la partie I

- Aben, Job, Greta Bocedi, Stephen C. F. Palmer, Petri Pellikka, Diederik Strubbe, Caspar Hallmann, Justin M. J. Travis, Luc Lens et Erik Matthysen (08/2016). “The importance of realistic dispersal models in conservation planning : application of a novel modelling platform to evaluate management scenarios in an Afrotropical biodiversity hotspot”. en. In : *Journal of Applied Ecology* 53.4, p. 1055-1065. URL : <http://onlinelibrary.wiley.com/doi/10.1111/1365-2664.12643/abstract> (cf. pp. 5, 59).
- Allema, Bas, Wopke van der Werf, Joop C. van Lenteren, Lia Hemerik et Walter A. H. Rossing (12/2014). “Movement Behaviour of the Carabid Beetle *Pterostichus melanarius* in Crops and at a Habitat Interface Explains Patterns of Population Redistribution in the Field”. In : *PLoS ONE* 9.12, e115751. URL : <http://dx.doi.org/10.1371/journal.pone.0115751> (cf. p. 9).
- Aviron, S., F. Burel, J. Baudry et N. Schermann (06/2005). “Carabid assemblages in agricultural landscapes : impacts of habitat features, landscape context at different spatial scales and farming intensity”. In : *Agriculture, Ecosystems & Environment*. Agri-Environmental Schemes as Landscape Experiments Agri-Environmental Schemes as Landscape Experiments 108.3, p. 205-217. URL : <http://www.sciencedirect.com/science/article/pii/S0167880905000514> (cf. p. 7).
- Bourhis, Yoann, Sylvain Poggi, Youcef Mammeri, Ronan Le Cointe, Anne Marie Cortesero et Nicolas Parisey (03/2017). “Foraging as the landscape grip for population dynamics—A mechanistic model applied to crop protection”. In : *Ecological Modelling* 354, p. 26-36 (cf. p. 5).
- Burel, Françoise et Jacques Baudry (11/1999). *Ecologie du paysage. Concepts, méthodes et applications*. Français. Paris : Tec & Doc Lavoisier (cf. p. 2).
- Chaplin-Kramer, Rebecca, Megan E. O'Rourke, Eleanor J. Blitzer et Claire Kremen (09/2011). “A meta-analysis of crop pest and natural enemy response to landscape complexity”. en. In : *Ecology Letters* 14.9, p. 922-932. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2011.01642.x/abstract> (cf. p. 7).
- Duelli, Peter (04/1997). “Biodiversity evaluation in agricultural landscapes : An approach at two different scales”. In : *Agriculture, Ecosystems & Environment*. Biodiversity in Agriculture - for a Sustainable Future 62.2, p. 81-91. URL : <http://www.sciencedirect.com/science/article/pii/S0167880996011437> (cf. pp. 2, 25).
- Dunning, John B., Brent J. Danielson et H. Ronald Pulliam (1992). “Ecological Processes That Affect Populations in Complex Landscapes”. In : *Oikos* 65.1, p. 169-175. URL : <http://www.jstor.org/stable/3544901> (cf. p. 4).
- Eyre, M. D., M. L. Luff et C. Leifert (01/2013). “Crop, field boundary, productivity and disturbance influences on ground beetles (Coleoptera, Carabidae) in the agroecosystem”. In : *Agriculture, Ecosystems & Environment* 165, p. 60-67. URL : <http://www.sciencedirect.com/science/article/pii/S0167880912004549> (cf. p. 7).
- Fahrig, Lenore, Jacques Baudry, Lluís Brotons, Françoise G. Burel, Thomas O. Crist, Robert J. Fuller, Clelia Sirami, Gavin M. Siriwardena et Jean-Louis Martin (02/2011). “Functional landscape heterogeneity and animal biodiversity in agricultural landscapes”. en. In : *Ecology Letters* 14.2, p. 101-112. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2010.01559.x/abstract> (cf. pp. 2, 26, 59, 64).

-
- Fahrig, Lenore, Judith Girard, Dennis Duro, Jon Pasher, Adam Smith, Steve Javorek, Douglas King, Kathryn Freemark Lindsay, Scott Mitchell et Lutz Tischendorf (02/2015). “Farmlands with smaller crop fields have higher within-field biodiversity”. In : *Agriculture, Ecosystems & Environment* 200, p. 219-234. URL : <http://www.sciencedirect.com/science/article/pii/S0167880914005192> (cf. pp. 7, 9, 26, 43, 59, 63).
- Forman, Richard T. T. (11/1995a). *Land Mosaics : The Ecology of Landscapes and Regions*. en. Google-Books-ID : sSRNU_5P5nwC. Cambridge University Press (cf. p. 2).
- (06/1995b). “Some general principles of landscape and regional ecology”. en. In : *Landscape Ecology* 10.3, p. 133-142. URL : <https://link.springer.com/article/10.1007/BF00133027> (cf. p. 2).
- Forman, Richard T. T. et Michel Godron (02/1986). *Landscape Ecology*. en. Google-Books-ID : ZvNEVs2MWqcC. Wiley (cf. p. 2).
- Fortin, Marie-Josée et Mark RT Dale (2005). *Spatial analysis : a guide for ecologists*. Cambridge University Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=bTT7tEiTFLoC&oi=fnd&pg=PA1&dq=spatial+analysis:+a+guide+for+ecologists.&ots=qSgTFacVRj&sig=gAXaISwd8Y86GLbd0tn3z67T7vA> (cf. pp. 4, 60).
- Gardiner, M. M., D. A. Landis, C. Gratton, C. D. DiFonzo, M. O’Neal, J. M. Chacon, M. T. Wayo, N. P. Schmidt, E. E. Mueller et G. E. Heimpel (01/2009). “Landscape diversity enhances biological control of an introduced crop pest in the north-central USA”. In : *Ecological Applications* 19.1, p. 143-154. URL : <http://www.esajournals.org/doi/abs/10.1890/07-1265.1> (cf. p. 7).
- Grimm, Volker, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M. Mooij, Steven F. Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand et Donald L. DeAngelis (11/2005). “Pattern-Oriented Modeling of Agent-Based Complex Systems : Lessons from Ecology”. en. In : *Science* 310.5750, p. 987-991. URL : <http://www.sciencemag.org/content/310/5750/987> (cf. pp. 8-10, 56, 60, 96, 100).
- Holling, C. S. et Craig R. Allen (06/2002). “Adaptive Inference for Distinguishing Credible from Incredible Patterns in Nature”. en. In : *Ecosystems* 5.4, p. 319-328. URL : <https://link.springer.com/article/10.1007/s10021-001-0076-2> (cf. pp. 9, 35, 96).
- Jongejans, Eelke, Olav Skarpaas et Katriona Shea (03/2008). “Dispersal, demography and spatial population models for conservation and control management”. In : *Perspectives in Plant Ecology, Evolution and Systematics*. Space matters - Novel developments in plant ecology through spatial modelling 9.3-4, p. 153-170. URL : <http://www.sciencedirect.com/science/article/pii/S1433831907000467> (cf. pp. 5, 8, 59, 61, 81).
- Jonsson, Mattias, Riccardo Bommarco, Barbara Ekbom, Henrik G. Smith, Jan Bengtsson, Berta Caballero-Lopez, Camilla Winqvist et Ola Olsson (03/2014). “Ecological production functions for biological control services in agricultural landscapes”. en. In : *Methods in Ecology and Evolution* 5.3, p. 243-252. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12149/abstract> (cf. pp. 5, 59).
- Landis, Douglas A., Stephen D. Wratten et Geoff M. Gurr (2000). “Habitat Management to Conserve Natural Enemies of Arthropod Pests in Agriculture”. In : *Annual Review of En-*

-
- tomology* 45.1, p. 175-201. URL : <http://dx.doi.org/10.1146/annurev.ento.45.1.175> (cf. pp. 5, 24, 59).
- Marrec, Ronan, Isabelle Badenhauer, Vincent Bretagnolle, Luca Börger, Marilyn Roncoroni, Nadine Guillon et Bertrand Gauffre (01/2015). “Crop succession and habitat preferences drive the distribution and abundance of carabid beetles in an agricultural landscape”. In : *Agriculture, Ecosystems & Environment* 199, p. 282-289. URL : <http://www.sciencedirect.com/science/article/pii/S016788091400468X> (cf. pp. 7, 9, 24, 33, 67).
- McIntire, Eliot J. B. et Alex Fajardo (01/2009). “Beyond description : the active and effective way to infer processes from spatial patterns”. In : *Ecology* 90.1, p. 46-56. URL : <http://www.esajournals.org/doi/abs/10.1890/07-2096.1> (cf. pp. 8, 9, 54, 56, 60).
- McIntire, Eliot J. B., Ghislain Rompré et Paul M. Severns (05/2013). “Biased correlated random walk and foray loop : which movement hypothesis drives a butterfly metapopulation?” en. In : *Oecologia* 172.1, p. 293-305. URL : <http://link.springer.com/article/10.1007/s00442-012-2475-9> (cf. pp. 10, 54, 60, 100).
- Moloney, Kirk A. et Simon A. Levin (1996). “The effects of disturbance architecture on landscape-level population dynamics”. In : *Ecology*, p. 375-394. URL : <http://www.jstor.org/stable/2265616> (cf. pp. 5, 59).
- Nathan, Ran (09/2008). “An emerging movement ecology paradigm”. en. In : *Proceedings of the National Academy of Sciences* 105.49, p. 19050-19051. URL : <http://www.pnas.org/content/105/49/19050> (cf. pp. 10, 54, 60).
- Parisey, Nicolas, Yoann Bourhis, Lionel Roques, Samuel Soubeyrand, Benoît Ricci et Sylvain Poggi (12/2016). “Rearranging agricultural landscapes towards habitat quality optimisation : In silico application to pest regulation”. In : *Ecological Complexity* 28, p. 113-122. URL : <https://www.sciencedirect.com/science/article/pii/S1476945X16300538> (cf. pp. 5, 34, 59, 101).
- Peck, Steven L. (10/2004). “Simulation as experiment : a philosophical reassessment for biological modeling”. eng. In : *Trends in Ecology & Evolution* 19.10, p. 530-534 (cf. pp. 8-10, 96, 100).
- Puech, Camille, Sylvain Poggi, Jacques Baudry Jacques et Stéphanie Aviron Stéphanie (10/2014). “Do farming practices affect natural enemies at the landscape scale?” In : *Landscape Ecology* 30 (cf. p. 7).
- Ricci, Benoît, Pierre Franck, Muriel Valantin-Morison, David A. Bohan et Claire Lavigne (09/2013). “Do species population parameters and landscape characteristics affect the relationship between local population abundance and surrounding habitat amount?” In : *Ecological Complexity* 15, p. 62-70. URL : <http://www.sciencedirect.com/science/article/pii/S1476945X13000330> (cf. p. 8).
- Savage, David et Michael Renton (01/2014). “Requirements, design and implementation of a general model of biological invasion”. In : *Ecological Modelling* 272, p. 394-409. URL : <http://www.sciencedirect.com/science/article/pii/S0304380013004638> (cf. pp. 5, 59).
- Schröder, B. et R. Seppelt (12/2006). “Analysis of pattern–process interactions based on landscape models—Overview, general concepts, and methodological issues”. In : *Ecological Modelling. Pattern and Processes of Dynamic Mosaic Landscapes – Modelling, Simulation, and Implications* 199.4, p. 505-516. URL : <http://www.sciencedirect.com/science/article/pii/S0304380006002936> (cf. pp. 3, 4, 59-61, 96).

-
- Schweiger, Andreas H., Severin D. H. Irl, Manuel J. Steinbauer, Jürgen Dengler et Carl Beierkuhnlein (04/2016). “Optimizing sampling approaches along ecological gradients”. en. In : *Methods in Ecology and Evolution* 7.4, p. 463-471. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12495/abstract> (cf. pp. 10, 60, 101).
- Tilman, David et Peter M. Kareiva (1997). *Spatial ecology : the role of space in population dynamics and interspecific interactions*. T. 30. Princeton University Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=QdqC8PvWB7AC&oi=fnd&pg=PR7&dq=Spatial+ecology:+the+role+of+space+in+population+dynamics+and+interspecific+interactions.&ots=bKYtHhZj&sig=4VpUdp00mt23W37oQZ6aRi7BJAY> (cf. pp. 8, 60).
- Tuda, Midori (03/2007). “Understanding mechanism of spatial ecological phenomena : a preface to the special feature on “Spatial statistics””. en. In : *Ecological Research* 22.2, p. 183-184. URL : <http://link.springer.com/article/10.1007/s11284-006-0328-x> (cf. pp. 8, 60).
- Turner, Monica G. et Robert H. Gardner (2015). *Landscape Ecology in Theory and Practice*. en. New York, NY : Springer New York. URL : <http://link.springer.com/10.1007/978-1-4939-2794-4> (cf. pp. 2, 6-8).
- Vasseur, Chloé, Alexandre Joannon, Stéphanie Aviron, Françoise Burel, Jean-Marc Meynard et Jacques Baudry (02/2013). “The cropping systems mosaic : How does the hidden heterogeneity of agricultural landscapes drive arthropod populations?” In : *Agriculture, Ecosystems & Environment*. Landscape ecology and biodiversity in agricultural landscapes 166, p. 3-14. URL : <http://www.sciencedirect.com/science/article/pii/S0167880912003246> (cf. pp. 2, 25).
- Vinatier, Fabrice (01/2010). *Dynamique spatiale du charançon du bananier en interaction avec le système de culture et l'organisation paysagère*. AgroParisTech. URL : <http://www.theses.fr/2010AGPT0068> (cf. p. 4).
- Vinatier, Fabrice, Philippe Tixier, Pierre-François Duyck et Françoise Lescourret (2011). “Factors and mechanisms explaining spatial heterogeneity : a review of methods for insect populations”. en. In : *Methods in Ecology and Evolution* 2.1, p. 11-22. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2010.00059.x/abstract> (cf. pp. 9, 81, 100).
- Wiens, John A. (1999). “The Science and Practice of Landscape Ecology”. en. In : *Landscape Ecological Analysis*. DOI : 10.1007/978-1-4612-0529-6_16. Springer, New York, NY, p. 371-383. URL : https://link.springer.com/chapter/10.1007/978-1-4612-0529-6_16 (cf. p. 3).
- Winsberg, Eric (01/2003). “Simulated Experiments : Methodology for a Virtual World”. In : *Philosophy of Science* 70.1, p. 105-125. URL : <http://www.journals.uchicago.edu/doi/abs/10.1086/367872> (cf. p. 10).
- Zurell, Damaris *et al.* (04/2010). “The virtual ecologist approach : simulating data and observers”. en. In : *Oikos* 119.4, p. 622-635. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0706.2009.18284.x/abstract> (cf. pp. 10, 54, 60).



Deuxième partie

Étude de l'effet de l'hétérogénéité de configuration du paysage sur les dynamiques de population dans un contexte de connaissance limitée sur les processus écologiques

1	Présentation de la démarche, du modèle biologique et de la variable paysagère	20
1.1	Démarche méthodologique	20
1.2	Présentation du modèle biologique	23
1.3	L'hétérogénéité de configuration du paysage : définition, pertinence et métriques	25
2	Matériel et Méthode	28
2.1	Génération de paysages agricoles virtuels	28
2.2	Typologie des processus en fonction des motifs	33
3	Résultats et discussion	40
3.1	Génération des paysages virtuels	40
3.2	Typologie des modèles de processus sur la base de leur signature populationnelle	41
3.3	Relation entre les métriques de population et la variable paysagère . .	43
4	Conclusion du chapitre 2	47
	Références	52

1

Présentation de la démarche, du modèle biologique et de la variable paysagère

1.1 Démarche méthodologique

Comme nous l'avons vu précédemment, l'aller-retour entre observations sur le terrain et modélisation permet d'améliorer de manière robuste la connaissance du lien entre processus écologiques et motifs de population (cf. figure 1.4) :

- L'étape initiale d'observation des motifs permet de formuler des hypothèses sur les processus sous-jacents
- Les hypothèses en compétition sont traduites en modèles mécanistes, qui sont simulés sur un espace explicite
- Les motifs générés sont comparés avec les observations de terrain. Les hypothèses les moins plausibles concernant les processus sont écartées, d'autres hypothèses peuvent être affinées ou de nouvelles hypothèses peuvent être formulées en regard de l'adéquation des motifs avec les observations sur le terrain.

On réitère ce processus jusqu'à ce qu'émerge une hypothèse robuste sur les processus écologiques. A partir de là, des scénarios d'aménagement du paysage peuvent être explorés *in silico* et il est possible d'en déduire des aménagements favorables en regard de la gestion de l'espèce ou du peuplement d'intérêt.

Si elle permet d'assurer la robustesse des relations entre processus et motifs, et *a fortiori* des conclusions associées à l'exploration de scénarios d'aménagement du paysage, cette démarche est fastidieuse, impliquant des observations sur le terrain et les coûts humains et logistiques associés . De plus, l'échelle temporelle associée à cette démarche semble être incompatible avec le besoin pressant d'aménagement du paysage pour la conservation d'espèces (Hallmann *et al.* 2017) ou la provision de services écosystémiques (Bommarco *et al.* 2013). Nous avons étudié si cette démarche peut-être court-circuitée, *i.e.* si les connaissances à notre disposition sur les processus écologiques régissant les dynamiques de population d'un organisme d'intérêt sont suffisantes pour suggérer des préconisations en termes d'aménagement du paysage dans une optique de gestion de population.

En comparant les motifs issus des différents processus écologiques possibles selon différents scénarios d'aménagement du paysage, deux types de résultats peuvent émerger :

- L'ensemble des processus génèrent des motifs de population similaires quel que soit l'aménagement paysager, *i.e.* les descripteurs de population issus des différents processus covarient avec la variable paysagère.
- Tous les processus ne génèrent pas des motifs similaires, les hypothèses ne réagissent pas toutes de la même manière à la variable paysagère.

Dans le premier cas, il ne sera pas nécessaire d'approfondir les connaissances sur les processus puisque tous réagissent de la même manière à l'aménagement du paysage. La valeur de la variable paysagère qui sera sélectionnée sera celle qui répondra au mieux aux objectifs de gestion, *i.e.* qui produira des motifs favorables quel que soit le processus sous-jacent. Dans le deuxième cas, il existera plusieurs types de réponses à la variable paysagère selon la nature des processus. Un échantillonnage sur le terrain s'avèrera nécessaire pour inférer les processus à partir de la comparaison des motifs simulés et des observations, mais cette étape aura au moins permis de réduire le nombre d'hypothèses en compétition en rassemblant les processus générant des réponses similaires à la variable paysagère.

La comparaison des motifs de population issus de différentes hypothèses écologiques a donc deux issues favorables possibles : *i)* dans le cas où les motifs sont similaires, elle permet l'économie d'une campagne d'acquisition de données d'observation supplémentaire sur le terrain ou *ii)* elle facilite l'inférence des processus écologiques en réduisant le nombre d'hypothèses en compétition. La mise en place d'une méthode qui détermine *a priori* si les connaissances sur les processus écologiques sont suffisantes pour statuer sur l'effet de la variable paysagère sur les motifs de population apparaît donc comme cruciale dans la recherche de paysages favorables d'un organisme d'intérêt.

Dans ce chapitre, nous présentons une méthode heuristique permettant de répondre à ces objectifs, et illustrons cette démarche dans le contexte général de la recherche de paysages agricoles favorables à l'amélioration des services écosystémiques de régulation biologique. Nous appliquons cette méthode heuristique à un coléoptère carabique, *Pterostichus melanarius*, dont les motifs de population à l'échelle du paysage agricole ont été largement décrits (e.g. [Holland, Thomas et al. 2005](#)), mais pour lequel les connaissances sur les processus ayant générés ces motifs sont encore insuffisantes pour permettre la recherche de paysages favorables via une exploration *in silico*. Nous déterminons si la diversité des hypothèses existantes sur les processus

écologiques est un obstacle pour caractériser le lien entre les motifs de populations et une métrique d'hétérogénéité de configuration du paysage, particulièrement pertinente dans les agro-écosystèmes.

Notre démarche méthodologique consiste en quatre étapes :

1. Génération de paysages présentant un gradient d'une variable paysagère d'intérêt.

Dans notre cas d'étude, la variable paysagère d'intérêt est une métrique d'hétérogénéité de configuration du paysage, la densité d'interfaces (Edge density ED).

2. Formulation de modèles de dynamiques de population porteurs des hypothèses écologiques plausibles régissant la dynamique au sein du paysage de l'organisme d'intérêt.

Dans notre cas d'étude, le formalisme des modèles de dynamique de population est un formalisme de réaction-diffusion. Les hypothèses écologiques diffèrent dans la dépendance de deux traits de vie (démographie et dispersion) à l'hétérogénéité spatiale des ressources (types de couvert). Les modèles formulés pour représenter les différentes hypothèses sont appelés Modèles de Processus (PM).

3. Simulation dynamique et spatialement explicite des modèles de dynamiques de populations sur les paysages générés, et extraction des motifs de population associés à chaque modèle.

Dans notre cas d'étude, les motifs de population sont des statistiques résumées des dynamiques de population agrégées, extraites à l'échelle du paysage et des différents types d'occupation du sol.

4. Typologie des modèles selon leur réponse à la variable paysagère.

Dans notre cas d'étude, les modèles de processus sont regroupés, via un algorithme de classification, sur la base de leurs motifs pour chaque valeur d'ED. Les processus qui sont regroupés quelle que soit la valeur d'ED constituent ce que nous appellerons dans ce chapitre un "écotype".

5. Comparaison des motifs de population associés à chaque écotype.

L'objectif est double. D'une part, comprendre sur la base de quels traits de vie (démographie ou mouvement) les écotypes sont formés permet de hiérarchiser leur importance sur la formation des motifs. D'autre part, rassembler les modèles selon leur réponse à la variable paysagère d'intérêt permet de réduire le nombre initial

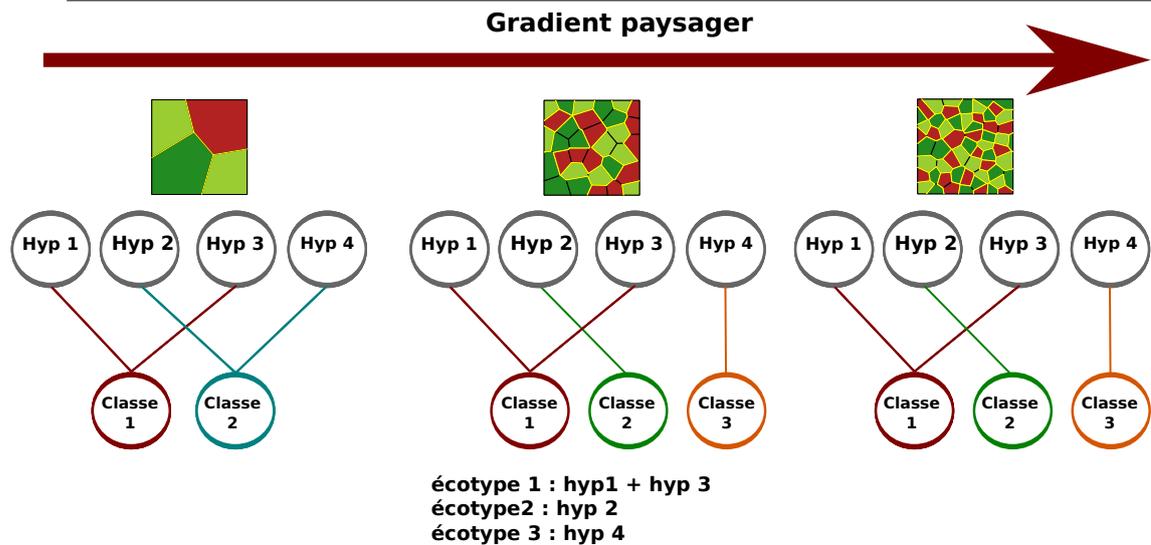


FIGURE 1.1 – Regroupement des modèles selon la similarité de leurs signatures populationnelles le long d’un gradient paysager. Lorsque des hypothèses écologiques amènent à des signatures populationnelles similaires, on peut considérer que les mécanismes cruciaux ont été capturés. Il n’est pas besoin d’approfondir les connaissances écologiques pour tester l’effet de la variable paysagère sur la dynamique de population. Dans le cas où les signatures sont différentes, cependant, de nouvelles observations sur le terrain seront nécessaires pour statuer sur les mécanismes responsables des dynamiques de population sur le terrain.

d’hypothèses en compétition sur les processus, et donc de faciliter l’inférence des processus à partir des motifs (figure 1.1).

Les deux sections qui suivent affinent la présentation du cas d’étude, en exposant le modèle biologique et précisant la métrique d’hétérogénéité de configuration du paysage. La deuxième partie décrit le modèle permettant de générer des paysages présentant un gradient de la variable paysagère, la manière dont ses paramètres ont été bornés à partir de l’analyse d’un paysage agricole réel et les résultats concernant la génération des paysages. La troisième partie présente la manière dont les hypothèses sur les processus écologiques ont été formulées en modèles de réaction-diffusion, la simulation de ces modèles sur les paysages générés, la classification des modèles sur la base de leurs motifs de population. Nous présenterons les résultats, et discuterons de l’implication et des limites de notre méthode dans cette dernière partie.

1.2 Présentation du modèle biologique

Les représentants de la famille des carabidés sont riches en diversité et en abondance au sein des agro-écosystèmes de tous les continents. Cette ubiquité, additionnée à la diversité de leurs régimes alimentaires (entomophages, granivores, malacophages ou polyphages opportunistes) et la voracité de certaines espèces font que

les carabidés sont considérés comme ayant un potentiel important de régulation des ravageurs au sein des agro-écosystèmes (Kromp 1999 ; Holland J. M 2004). L'aménagement du paysage agricole pour favoriser les carabes est vu comme une alternative prometteuse à la lutte chimique pour la régulation des bio-agresseurs (Landis *et al.* 2000 ; Bommarco *et al.* 2013 ; Geiger *et al.* 2010), et les études pour mettre en lien la structure paysagère et les dynamiques de population de carabes se sont multipliées durant les 30 dernières années.

Au sein de cette famille, *Pterostichus melanarius* a reçu une attention particulière, étant l'une des 8 espèces les plus abondantes au sein des paysages agricoles d'Europe de l'Ouest (Kromp 1999). Son intérêt agronomique provient de son régime alimentaire à 90% carnivore, constitué d'une large gamme d'invertébrés épigés, incluant au moins 14 espèces de ravageurs appartenant aux mollusques, coléoptères, diptères et lépidoptères (Sunderland 2002). Le rôle de *P. melanarius* en tant que régulateur de ravageurs de culture a été démontré pour les limaces et les pucerons (Winder *et al.* 2005).

Les adultes émergent des pupes vers la fin du printemps, avec un pic en juin-juillet, et plutôt des parcelles cultivées que des habitats semi-naturels (Wallin et Ekblom 1988 ; Marrec *et al.* 2015). *P. melanarius* se déplace majoritairement en marchant plutôt qu'en volant. Si ces connaissances sur la biologie de *P. melanarius* font consensus, les informations fines sur les valeurs des traits de vie et leur hétérogénéité au sein du paysage agricole sont plus disparates et parfois contradictoires. Selon les études, *P. melanarius* n'est inféodé à aucun type de culture en particulier (Kromp 1999) ou bien est plutôt inféodé aux champs de céréales (Thiele 1977), d'autres études montrent que les préférences d'habitat changent au cours de la saison, du fait de l'asynchronie des phénologies des cultures et des conditions microclimatiques associées (CÁrcamo et Spence 1994 ; Honek 1988).

Les informations sur les capacités de dispersion de *P. melanarius* au sein des différents éléments de la mosaïque paysagère sont elles aussi disparates. Si le mouvement au sein de parcelles de différents types culturels a été quantifié dans différentes études (*e.g.* Wallin et Ekblom 1988 ; Wallin et Ekblom 1994 ; Thomas *et al.* 2006), la diversité des méthodes employées ne permet pas de comparer les résultats. Au niveau des interfaces, des différences de flux nets entre différents couverts ont été observés, mais il n'a pas été possible de déterminer si ce phénomène relevait d'une simple différence dans les vitesses de déplacement de part et d'autre de l'interface, ou d'un processus actif de mouvement dirigé de l'habitat le moins favorable vers le plus favorable (Baars 1979 ; Rijnsdorp 1980). En résumé, l'état actuel de la connaissance sur les distributions et les patterns de mouvement des prédateurs épigés est insuffisant pour fournir des préconisations sur la manière dont le paysage agricole

doit être arrangé pour bénéficier au maximum des services de régulation fournis par *P. melanarius* (Alema 2014). Cependant, savoir dans quelle mesure l'acquisition ou non de ces connaissances fines sur les processus est nécessaire pour préconiser des aménagements de paysages favorables est également crucial.

1.3 L'hétérogénéité de configuration du paysage : définition, pertinence et métriques

Depuis les prémisses de l'écologie du paysage et jusqu'à relativement récemment, le paysage agricole était traditionnellement vu comme un paysage dual, constitué d'îlots d'habitats favorables (les habitats semi-naturels comme les prairies ou les haies), séparés par une matrice hostile qui comprenait tous les autres types d'occupation du sol.

Les couverts agricoles, cependant, sont pourvoyeurs de ressources, au moins temporairement, et supportent une biodiversité elle-même potentiellement pourvoyeuse de services écosystémiques (Vasseur *et al.* 2013). Duelli (1997), prenant en compte le fait que les couverts agricoles puissent être des habitats favorables, a pour la première fois proposé le concept de mosaïque paysagère. Ce changement de paradigme a mené à un changement dans la manière de décrire le paysage et a fait émerger les concepts d'hétérogénéités de composition et de configuration du paysage (Fahrig et Nuttle 2005). L'hétérogénéité de composition correspond à la diversité et à la quantité des types d'occupation du sol constituant le paysage, et l'hétérogénéité de configuration correspond à la manière dont ceux-ci sont agencés dans l'espace (cf. figure 1.2).

Dans un contexte agricole, où les types cultureux sont contraints par les besoins de production, la marge de manœuvre sur l'hétérogénéité de composition peut être limitée à l'échelle du paysage. Au-delà de cet argument agronomique, Fahrig et Nuttle (2005) posent plusieurs hypothèses pour encourager l'étude de l'effet de l'hétérogénéité de configuration sur les motifs de population dans le paysage agricole. Ces hypothèses posent le mouvement entre patches comme un mécanisme crucial. Premièrement, à composition égale, un paysage présentant une hétérogénéité de configuration plus élevée présentera des patches de plus petite taille, présentant un ratio périmètre/aire plus élevé, et une distance entre centroïdes plus faible. Les longueurs d'interfaces entre couverts différents seront également plus élevées. En somme, les échanges entre patches sont facilités lorsque l'hétérogénéité de configuration est plus élevée. Cela facilite les processus de complémentation et de supplémentation définis dans le chapitre 1 : un organisme qui a besoin de ressources fournies par des couverts différents parviendra plus facilement à y accéder dans un paysage hétérogène. Dans

le cas des couverts agricoles, où les temporalités des qualités d'habitat associées aux phénologies des différents types culturaux sont asynchrones, le mouvement des individus d'un patch où les ressources déclinent vers un patch où les ressources sont disponibles doit avoir un effet positif sur la population.

Dans un travail plus empirique, [Fahrig, Girard *et al.* \(2015\)](#) montrent ainsi que l'augmentation de l'hétérogénéité de configuration, indépendamment de l'hétérogénéité de composition, est favorable à la biodiversité.

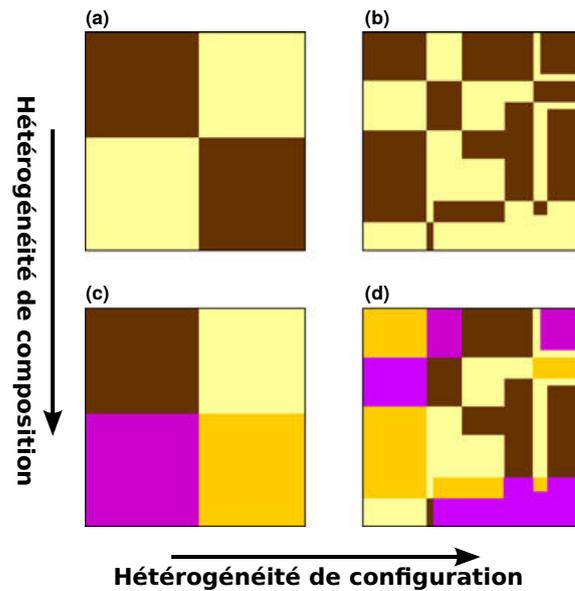


FIGURE 1.2 – Illustration des deux axes majeurs de l'hétérogénéité spatiale : hétérogénéité de configuration et hétérogénéité de composition. Chaque carré est un paysage, les couleurs représentent les différents types d'occupations du sol. Adapté de [Fahrig, Baudry *et al.* \(2011\)](#).

Les mesures d'hétérogénéité de configuration incluent des métriques comme la taille moyenne des patches, la densité d'interfaces ou des indices d'interspersión/juxtaposition ([Cushman *et al.* 2008](#) ; [Fahrig, Baudry *et al.* 2011](#)).

En gardant constante la proportion des différents types d'occupation du sol, la densité d'interfaces entre couverts différents (Edge density ED) est un bon proxy vers l'hétérogénéité de configuration, puisqu'il conditionne à la fois la taille des patches (et donc la distance entre centroïdes) et l'interspersión des différents patches ([Cushman *et al.* 2008](#) ; [Fahrig et Nuttle 2005](#)). Elle est définie comme suit :

$$ED = \frac{\sum_{i=1}^M E_{ik}}{A} \quad (1.1)$$

avec M nombre de types de patches, E_{ik} somme des longueurs d'interfaces entre type de patch i et type de patch k , et A surface du paysage.

Dans la partie suivante, nous décrivons le modèle permettant de générer des paysages présentant un gradient de cette métrique d'hétérogénéité de configuration.

2

Matériel et Méthode

2.1 Génération de paysages agricoles virtuels

L'objectif ici est de générer le support virtuel, le substrat pour la modélisation des dynamiques de population, *i.e.* le paysage. Ce support doit présenter des propriétés structurales, *i.e.* des indices paysagers (Cushman *et al.* 2008 ; McGarigal 2006), que l'on doit pouvoir contrôler pour tester les hypothèses écologiques. On fait appel à la modélisation de paysages pour pouvoir générer de multiples instances de paysages dont les propriétés structurales sont contrôlées (Gaucherel 2008).

2.1.1 Le modèle de paysage

Le paysage agricole est un paysage fortement anthropisé présentant une structure constituée de patches, qui sont des unités que l'on considérera homogènes, avec des interfaces nettes, bien délimitées (Lausch *et al.* 2015). Cette structure en patches peut être modélisée par la tessellation de processus de points marqués, qui génère un espace constitué d'unités non chevauchantes, séparées par des interfaces linéaires (Gaucherel 2008 ; Le Ber *et al.* 2009).

Définition d'une tessellation

Une *tessellation* divise le plan en polygones non-chevauchants ou l'espace en polyèdres (Illian *et al.* 2008). C'est une structure secondaire, construite à partir d'un motif de points (les germes) généré par un processus ponctuel dans un plan. Le modèle le plus usité de tessellation est la tessellation de Voronoï. Ce modèle construit des tuiles sur la base des distances euclidiennes entre germes : chaque tuile $T(y)$ est constituée de l'ensemble des points de l'espace dont le germe le plus proche est y . Formellement, tout point x du plan (exceptés les sommets et les points situés aux interfaces) a un unique plus proche voisin $n(x)$. La tuile $T(y)$ d'un germe y est définie par :

$$T(y) = \{x \in \mathbb{R}^2 : n(x) = y\}$$

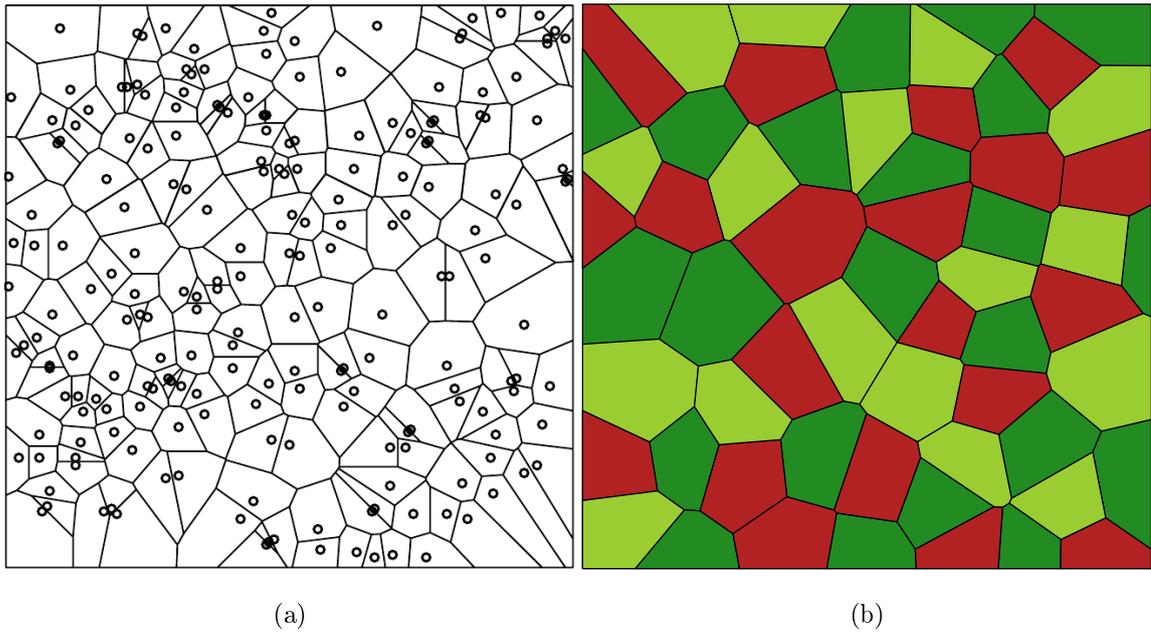


FIGURE 2.1 – Exemples de tessellations. (a) Issue d'un processus de points non marqués. Les zones de fortes densités de germes sont constituées de tuiles de plus faible surface. Extrait de [Illian et al. \(2008\)](#) (b) Issue d'un processus de points marqués.

La disposition des germes (noté par la suite le *processus de points*) définit donc la structure des polygones dans l'espace.

En ajustant les paramètres du processus de point, on peut modifier la taille, le nombre et l'arrangement des polygones. Il est également possible d'affecter des variables qualitatives (des marques) aux germes, définissant ainsi un processus de points marqués. Les polygones générés par la tessellation de tels processus de points peuvent hériter des attributs des germes.

On souhaite pouvoir contraindre dans notre modèle le nombre et la taille des polygones constituant les paysages virtuels pour leur donner des valeurs réalistes, *i.e.* rencontrées dans le paysage réel. La taille des polygones peut être traduite en une distance minimale tolérée entre centroïdes de polygones voisins. De plus, afin de centrer l'étude sur les effets de l'hétérogénéité de configuration du paysage, nous avons cherché à limiter la variabilité des tailles et des formes de patches au sein d'un paysage donné. Une famille de processus permettant de contraindre les distances entre centroïdes et donc la régularité des motifs des polygones est le modèle *Multitype Strauss-Hardcore (MSHC)*.

Il s'agit d'un processus de points marqué qui permet d'intégrer plusieurs types de marques (ici les types d'occupation du sol), des propriétés de distance *hardcore* (distance minimale entre marques) et des dépendances spatiales positives (attraction) ou négatives (répulsion) entre marques ([Edelman 2012](#)). En ajustant les paramètres du modèle MSHC, *i.e.* le nombre moyen de points marqués par unité de surface,

l'intensité et la direction des interactions entre marques, il est possible de générer des réalisations contrastées de processus de points marqués (Edelman 2012). Les propriétés de ces processus de points sont héritées par la tessellation (Illian *et al.* 2008 ; Baddeley 2010 ; Edelman 2012).

Définition du problème d'optimisation

Afin de générer des paysages présentant les propriétés recherchées en termes d'hétérogénéité de configuration, on résout de manière itérative le problème d'optimisation suivant :

$$\min_{(\beta_k, h_{km}, h_{kk}, \gamma_{km}, \gamma_{kk}, r_{km}, r_{kk})} \left(\sqrt{(\widehat{ED}(\mathbf{X}(\beta_k, h_{km}, \gamma_{km}, r_{km})) - \overline{ED})^2 + \sum_{m=1}^M \sqrt{(\widehat{A}_m(\mathbf{X}(\beta_k, h_{km}, \gamma_{km}, r_{km})) - \overline{A}_m)^2}} \right) \quad (2.1)$$

Avec

M nombre de types d'occupation du sol (marques)

\mathbf{X} Réalisation du processus de point obtenue par le modèle de MHSC

h_{km} distance minimale entre les points marqués de types différents

h_{kk} distance minimale entre les points marqués de même type

γ_{km} Intensité et direction de l'interaction entre points marqués de types différents

γ_{kk} Intensité et direction de l'interaction entre points marqués de même type

r_{km} Rayon d'interaction entre points marqués de types différents

r_{kk} Rayon d'interaction entre points marqués de même type

\widehat{A}_m Surface occupée par le type d'occupation du sol m après tessellation

\overline{A}_m Valeur cible de la surface occupée par le type d'occupation du sol m après tessellation

\overline{ED} Valeur cible de densité d'interfaces

\widehat{ED} Valeur de densité d'interface dans le paysage généré

L'algorithme d'optimisation opère de la façon suivante :

1. une première instance de paysage est générée à partir de la tessellation issue du processus de points obtenu avec les valeurs initiales des paramètres
2. les valeurs de la métrique ED (eq. 1.1) et des surfaces des différents types sont extraits de cette tessellation, et la fonction de coût est calculée
3. les valeurs de paramètres sont ajustées. L'algorithme d'optimisation utilisé est un algorithme global, le *multi-level single linkage*, qui part de plusieurs points de l'espace des paramètres à chaque itération, résolvant en local le problème

de minimisation par un processus de recuit-simulé.

4. le processus est répété et à chaque itération les valeurs des paramètres sont actualisées pour minimiser la fonction de coût. Comme on ne connaît pas la valeur de la fonction de coût associée à un paysage optimal, le nombre d'itérations a été choisi suffisamment grand pour permettre l'acquisition de paysages répondant aux valeurs ciblées d'*ED* et de surfaces des polygones des différents types, et suffisamment petit pour limiter le temps de calcul. La sélection des meilleurs paysages parmi les paysages générés est effectuée *a posteriori* selon la valeur finale de la fonction de coût.

Ce modèle de paysage a été programmé à l'aide du logiciel *R*. Les packages *spatstat*, *deldir*, *sp* et *nlotpr* ont été respectivement utilisés pour *i*) la génération des processus de points marqués, *ii*) la tessellation de l'espace à partir des processus de points, *iii*) la mesure des longueurs d'interfaces et des aires des polygones et *iv*) l'algorithme d'optimisation.

2.1.2 Calibration du modèle de paysage à partir d'un paysage réel

Afin de déterminer les bornes des paramètres du modèle de paysage et des valeurs cibles d'hétérogénéité du paysage, nous avons analysé des données géomatiques issues d'un paysage réel. La zone d'étude est la Zone Atelier Armorique (ZAA), un paysage agricole d'environ 130 km² situé dans le Nord-Ouest de la France (origine à 48°29' N 1° 34'W). En 2011, l'année sélectionnée pour l'analyse de paysage, la surface dédiée à l'agriculture (*i.e.* les prairies (semi-)permanentes et les parcelles cultivées) représentait $\sim 75\%$ de la surface, le reste étant partagé entre des surfaces boisées (14%) et des surfaces imperméabilisées (routes, bâti, etc...). La production agricole de cette zone est associée à un système de polyculture-élevage, et est donc composée en majorité de prairies, de céréales et de maïs (respectivement 46.8%, 23.3% et 27.2% de la surface agricole). À eux trois, ces couverts représentent plus de 97 % de la surface agricole de la ZAA (cf. figure 2.2).

Nous avons d'abord extrait la distribution des surfaces des parcelles agricoles à l'échelle de la ZAA. Pour calculer les longueurs d'interfaces, nous avons ensuite extrait des échantillons de paysages à l'aide de fenêtres glissantes de 300m×300m sur l'ensemble de la ZAA. L'emprise de la fenêtre glissante a été choisie à partir de la surface explorée par *P. melanarius* durant sa période d'activité, qui varie entre 3 et 12 ha selon les études et le mode de calcul (Alema 2014). En utilisant un pas de translation de 60 m, 6017 fenêtres ont été extraites sur la ZAA. Nous n'avons conservé que les fenêtres qui ne contenaient que des surfaces agricoles et qui pré-

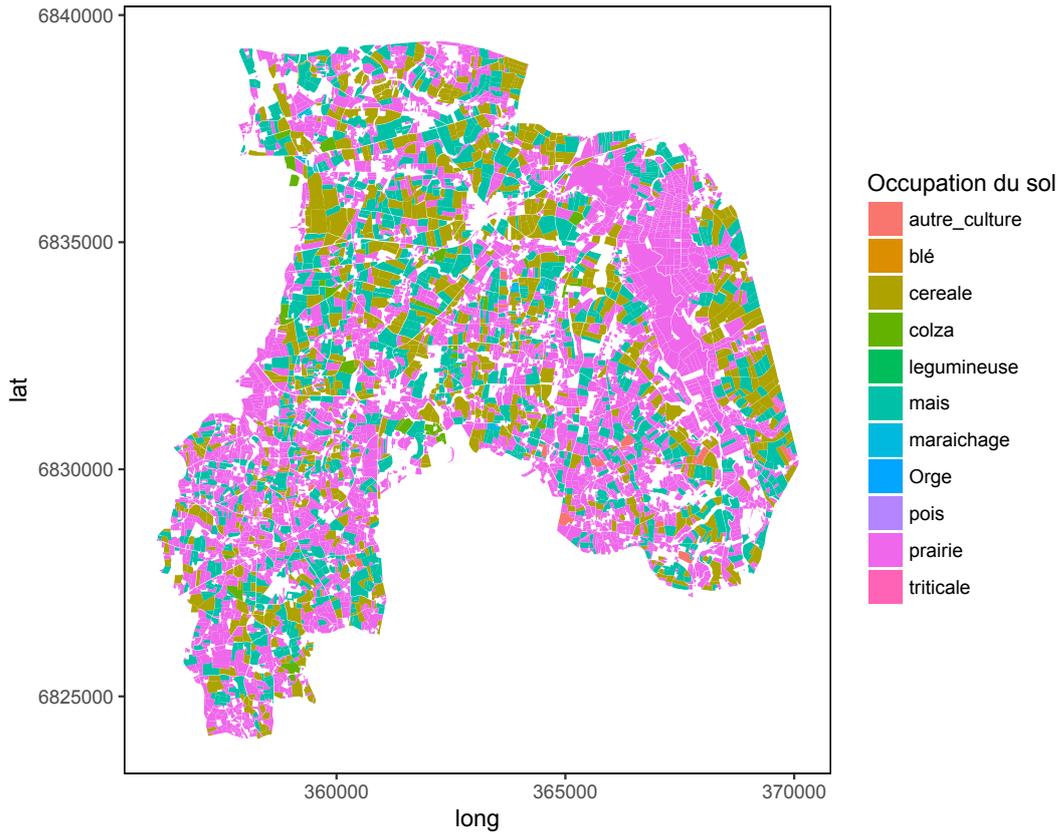


FIGURE 2.2 – Zone Atelier Armorique (parcelles agricoles uniquement, données 2011)

sentaient au moins une interface (*i.e.* qui comprenaient au moins 2 parcelles), soit 306 paysages. Nous avons ensuite extrait les longueurs d’interfaces au sein de ces paysages. À l’échelle de la ZAA, la taille des parcelles agricoles varie entre 0.1 et 70 ha. Les 5^{me} et 95^{me} percentiles donnent des valeurs de 0.2 et 4.2 hectares, respectivement. L’analyse des fenêtres glissantes donne des valeurs d’interfaces entre 0.58m/ha et 478,78 m/ha, ce qui exprimé par rapport à la longueur du côté de la fenêtre donne les valeurs sans unités de 0.02 et 14.36¹.

Traduction de l’analyse du paysage réel pour le modèle de paysage

Les valeurs mesurées de longueurs d’interfaces permettent de déterminer les valeurs cibles de la métrique ED pour le modèle de paysage. La taille minimale des parcelles à l’échelle du paysage permet de déterminer le paramètre β_k , qui est le nombre maximal de germes par unité de surfaces, *i.e.* le nombre maximal de polygones dans les paysages tessellés. Nous avons donc déterminé β comme le nombre maximal de parcelles que pourrait supporter le paysage si toutes les parcelles faisaient la surface minimale rencontrée dans le paysage réel, *i.e.* $\beta = 9/0.2 = 60$.

1. Exprimer les valeurs par rapport à la longueur du côté de la fenêtre permet de les comparer à l’unité de densité d’interface dans les paysages virtuels

Pour ne nous focaliser que sur l'effet de l'hétérogénéité de configuration, nous avons contraint l'équi-proportion surfacique des trois types d'occupation du sol au sein des paysages générés, *i.e.* $\overline{A_m} = 1/3$. Nous avons placé la limite inférieure des valeurs cibles d'ED à 3, considérant qu'elle est la valeur minimale possiblement atteinte au sein d'un paysage avec trois types de couverts en équiproportion. Nous avons élargi la valeur maximale d'ED à 29 (le double de la valeur réelle) pour explorer plus largement l'effet de l'hétérogénéité de configuration. Pour chaque valeur d'ED entre 3 et 29 par pas de 1, nous avons généré 60 instances de paysages. Les résultats sont présentés dans la section 3.1.

2.2 Typologie des processus en fonction des motifs

2.2.1 Description des modèles d'espèces

Comme nous l'avons dit précédemment, les connaissances sur les processus écologiques responsables des motifs de population de *P. melanarius* au sein des paysages agricoles sont largement méconnus, et l'état de la connaissance actuelle ne permet pas d'explorer des scénarios d'aménagement du paysage pour favoriser les services écosystémiques fournis par cette espèce (Alema 2014; Marrec *et al.* 2015). Notamment, les modalités de mouvement au sein des différents types de couverts agricoles sont peu connus, *i.e.* si le mouvement est homogène à l'échelle du paysage ou s'il est dépendant des types de couverts. De même ses préférences en termes d'habitat ne sont pas clairement établies, *i.e.* la démographie (balance natalité-mortalité) est-elle dépendante du type de couvert ?

Un modèle de réaction-diffusion est un modèle mathématique permettant de décrire au cours du temps un processus d'accroissement d'une grandeur continue (la réaction) et sa répartition dans l'espace (la diffusion). Formellement, ce modèle est représenté par des équations aux dérivées partielles. Le fait que ce formalisme soit basé sur des grandeurs continues et non sur des entités discrètes oblige le modélisateur en écologie à travailler à l'échelle de densités de populations, *i.e.* à faire une approximation des sommes des comportements individuels, et donc ne peut représenter des processus écologiques complexes intervenant à l'échelle individuelle ou interindividuelle. Cependant, ce formalisme permet de prendre en compte l'hétérogénéité du paysage en incluant des paramètres de démographie et/ou de diffusion dont les valeurs peuvent être fonction de l'espace et du temps, tout en étant parcimonieux et donc performants en termes de temps de calcul (Roques et Soubeyrand 2013). Ce compromis entre parcimonie et finesse des processus représentés fait que ce formalisme a été exploité dans différents domaines de l'écologie, incluant l'étude des invasions biologiques en milieu hétérogène (Shigesada et Kawasaki 1997; Roques,

Auger-Rozenberg *et al.* 2008) ou l'étude de l'effet de la structure paysagère sur les dynamiques de population d'insectes (Bourhis, Poggi, Mammeri, Cortesero *et al.* 2015; Parisey *et al.* 2016). Nous avons utilisé ce formalisme pour traduire mathématiquement les hypothèses sur les processus écologiques de *P. melanarius*.

La qualité d'habitat peut être représentée par le paramètre de capacité d'accueil, qui prend implicitement en compte des facteurs comme la quantité et la disponibilité des ressources ou les conditions (micro-)climatiques (Banks 1994; Barros *et al.* 2016). En lui donnant une valeur homogène, respectivement hétérogène, on peut représenter le caractère généraliste, respectivement spécialiste, de l'organisme d'intérêt (*e.g.* Padrón et Trevisan 2006). Nous avons donc modélisé la composante de réaction en utilisant le modèle de Verhulst (Banks 1994), qui permet de prendre en compte l'hétérogénéité spatiale en attribuant une valeur de capacité d'accueil dépendante, par exemple, du type d'occupation du sol.

La composante de diffusion a été modélisée par l'équation de Fokker-Planck pour les modèles à diffusion hétérogène, car elle permet de mieux représenter la redistribution des populations aux interfaces. Pour les modèles avec une diffusion homogène, la diffusion a été modélisée par l'équation de Fick (*i.e.* l'équation de diffusion de la chaleur) (Roques et Soubeyrand 2013). La formulation générale du modèle de réaction-diffusion s'écrit donc :

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K(\mathbf{x}, t)} \right) + \Delta_{\mathbf{x}} (\mu(\mathbf{x}, t)P(\mathbf{x}, t)) \quad (2.2)$$

avec \mathbf{x} la représentation de l'espace en deux dimensions (x, y) , $P(\mathbf{x}, t)$ la densité de population en \mathbf{x} au temps t , r le taux de croissance *per capita* de la population, $K(\mathbf{x}, t)$ la capacité d'accueil en \mathbf{x} au temps t , $\mu(\mathbf{x}, t)$ le paramètre de diffusion sur \mathbf{x} et t , appelé motilité quand il est appliqué à une diffusion de Fokker-Planck (Alema 2014). $\Delta_{\mathbf{x}} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ est l'opérateur Laplacien, représentant la dérivée seconde dans l'espace. Les formulations des modèles de processus diffèrent dans la manière dont les paramètres de motilité et de capacité d'accueil sont dépendants du type d'occupation du sol.

Afin de représenter la diversité des hypothèses sur les processus écologiques formulées pour expliquer les motifs de population observés empiriquement à l'échelle du paysage, nous avons pris le parti de formuler un grand nombre de modèles de processus. Parmi la gamme de modèles formulés, certains peuvent sembler trop simplistes pour être réalistes écologiquement. Cependant, si nous parvenons à montrer que des modèles simples parviennent à générer des motifs similaires à ceux générés par des modèles plus complexes, il peut être intéressant d'utiliser le modèle le

plus parcimonieux pour l'exploration de scénarios de paysages (Johnson et Omland 2004; Holling et Allen 2002). Nous avons donc défini six modalités pour chacune des composantes de réaction et de diffusion :

- Nulle *NULL* : Pas de changement de la densité de population à l'échelle du paysage ($r = 0$) ou de mouvement ($\mu = 0$)
- Homogène constante *HC* : La valeur du paramètre de capacité d'accueil (K) ou du paramètre de diffusion (μ) est homogène dans l'espace et ne varie pas au cours du temps
- Homogène décroissante *HD* : K ou μ est homogène dans l'espace et est une fonction décroissante exponentielle du temps.
- Homogène monovoltine *HV* : K ou μ est homogène dans l'espace varie dans le temps avec un pic à mi-saison (cf. détails en annexe)
- Inhomogène constante *IC* : K et/ou μ dépend du type d'occupation du sol et ne varie pas au cours du temps
- Inhomogène variant *IV* : K et/ou μ dépend du type d'occupation du sol et varie au cours du temps en suivant la qualité du couvert végétal, extrapolée à partir de valeurs de *Leaf Area Index* de trois types de couvert culturaux (cf. détails en annexe)

Nous avons effectué toutes les combinaisons possibles des modalités pour les composantes de réaction et de diffusion, puis avons écarté les modèles non-spatialisés, *i.e.* qui ne contenaient de dépendance à l'espace ni dans la composante de réaction ni de diffusion, donnant un total de 20 modèles spatialement explicites, formulés dans la table 2.1.

FIGURE 2.3 – Table 2 : Hypothèses écologiques pour les composantes de réaction et de diffusion, nom des modèles, formulation des modèles et paramètres impliqués. Les composantes de réaction et de diffusion diffèrent dans la manière dont la capacité d'accueil K et le paramètre de diffusion μ sont dépendant à l'espace et varient au cours du temps. HC : Homogène et Constant, HV : Homogène et Variant au cours du temps, IC : Hétérogène et Constant, IV : Hétérogène et Variant au cours du temps.

Reaction component	Diffusion component	Model name	Formulation $\frac{\partial P_{\mathbf{x},t}}{\partial t} =$	Parameters involved
NULL	IC	$\mathcal{M}_{Nr,ICd}$	$\frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x})P_{\mathbf{x},t})$	μ_1, μ_2, μ_3
NULL	IV	$\mathcal{M}_{Nr,IVd}$	$\frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x},t)P_{\mathbf{x},t})$	—
HC	IC	$\mathcal{M}_{HCr,ICd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x})P_{\mathbf{x},t})$	K, μ_1, μ_2, μ_3
HC	IV	$\mathcal{M}_{HCr,IVd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x},t)P_{\mathbf{x},t})$	K
HV	IC	$\mathcal{M}_{HVr,ICd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(t)}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x})P_{\mathbf{x},t})$	$r, K_m, \mu_1, \mu_2, \mu_3$
HV	IV	$\mathcal{M}_{HVr,IVd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(t)}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x},t)P_{\mathbf{x},t})$	K_m
IC	NULL	$\mathcal{M}_{ICr,Nd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x})}\right)$	r, K_1, K_2, K_3
IC	HC	$\mathcal{M}_{ICr,HCd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x})}\right) + \mu \left(\frac{\partial^2 P_{\mathbf{x},t}}{\partial \mathbf{x}^2} + \frac{\partial^2 P_{\mathbf{x},t}}{\partial y^2} \right)$	r, K_1, K_2, K_3, μ
IC	HV	$\mathcal{M}_{ICr,HVd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x})}\right) + \mu(t) \left(\frac{\partial^2 P_{\mathbf{x},t}}{\partial \mathbf{x}^2} + \frac{\partial^2 P_{\mathbf{x},t}}{\partial y^2} \right)$	r, K_1, K_2, K_3, μ_m
IC	IC	$\mathcal{M}_{ICr,ICd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x})}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x})P_{\mathbf{x},t})$	$r, K_1, K_2, K_3, \mu_1, \mu_2, \mu_3$
IC	IV	$\mathcal{M}_{ICr,IVd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x})}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x},t)P_{\mathbf{x},t})$	r, K_1, K_2, K_3
IV	NULL	$\mathcal{M}_{IVr,Nd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x},t)}\right)$	r
IV	HC	$\mathcal{M}_{IVr,HCd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x},t)}\right) + \mu \left(\frac{\partial^2 P_{\mathbf{x},t}}{\partial \mathbf{x}^2} + \frac{\partial^2 P_{\mathbf{x},t}}{\partial y^2} \right)$	r, μ
IV	HV	$\mathcal{M}_{IVr,HVd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x},t)}\right) + \mu(t) \left(\frac{\partial^2 P_{\mathbf{x},t}}{\partial \mathbf{x}^2} + \frac{\partial^2 P_{\mathbf{x},t}}{\partial y^2} \right)$	r, μ_m
IV	IC	$\mathcal{M}_{IVr,ICd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x},t)}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x})P_{\mathbf{x},t})$	r, μ_1, μ_2, μ_3
IV	IV	$\mathcal{M}_{IVr,IVd}$	$rP(\mathbf{x},t) \left(1 - \frac{P(\mathbf{x},t)}{K(\mathbf{x},t)}\right) + \frac{\partial^2}{\partial \mathbf{x}^2} (\mu(\mathbf{x},t)P_{\mathbf{x},t})$	r

2.2.2 Simulation des modèles

Les modèles formulés ont été simulés sur un échantillon des paysages générés. Afin de représenter le gradient d'hétérogénéité du paysage tout en limitant les temps de calcul et le volume de données, nous n'avons gardé que cinq valeurs d'ED (4, 9, 14, 19 et 24), et simulé les 20 modèles sur 15 instances de chacune de ces valeurs, soit 1500 simulations au total.

Pour la simulation numérique, l'espace et le temps ont été normalisés, *i.e.* nous avons considéré le domaine spatial comme un carré unitaire de 1×1 et la durée de simulation d'une unité de temps. Le domaine spatial Ω a été discrétisé et caractérisé par une matrice de valeurs dépendantes du type d'occupation du sol. La résolution spatiale est de 0.01×0.01 unités d'espace, définissant une grille de 100×100 pixels. Le temps est discrétisé en 1000 pas de temps. La simulation d'un modèle génère donc une matrice de $100 \times 100 \times 1000$.

La composante de réaction est résolue par un schéma explicite de Runge-Kutta 4 et la composante de diffusion est résolue à l'aide d'une méthode aux volumes fini et un schéma ADI (Alternate Direction Implicite) (Press 2007). Les conditions initiales de densités de population sont homogènes, placées à 0.5, *i.e.* la moitié de la capacité d'accueil maximale normalisée. Nous avons testé l'effet de conditions initiales aléatoires avec $P(\mathbf{x}, 0) \sim \mathcal{U}(0, 1)$ sur la classification des modèles de processus et en avons conclu que les résultats n'étaient pas différents de ceux obtenus avec des conditions homogènes. Les conditions au bord du domaine spatial sont des conditions réfléchissantes de Neumann, exprimées comme $\frac{\partial P(\mathbf{x}, t)}{\partial \mathbf{n}} \Big|_{\partial \Omega_i} = 0$ où \mathbf{n} est un vecteur unitaire normal à la frontière $\partial \Omega_i$.

2.2.3 Extraction des motifs de population

Pour chacun des PM simulés, nous avons extrait des statistiques résumées des populations virtuelles à l'échelle du paysage et de chacun des types d'occupation du sol. Chaque couple (PM, Ω) est ainsi décrit par un vecteur de huit descripteurs :

- Moyenne dans le temps de la moyenne de la population à l'échelle du paysage (Ω)

$$\overline{P_{\Omega}(t)} = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N P(\mathbf{x}_i, t), \mathbf{x} \in \Omega \quad (2.3)$$

- Moyenne dans le temps de l'écart-type de la population à l'échelle du paysage (Ω)

$$\sigma_{\Omega}(t) = \frac{1}{T} \sum_{t=1}^T \sqrt{\frac{1}{N} \sum_{i=1}^N (P(\mathbf{x}_i, t) - \overline{P(t)})^2}, \mathbf{x} \in \Omega \quad (2.4)$$

où $N = 10000$ le nombre d'unités de discrétisations dans Ω

- Moyenne dans le temps de la moyenne de la population à l'échelle des sous-domaines $\omega_1, \omega_2, \omega_3$ constitués par les 3 types d'occupation du sol

$$\overline{P_{\omega_j}(t)} = \frac{1}{N_j T} \sum_{t=1}^T \sum_{i=1}^{N_j} P(\mathbf{x}_i, t), \mathbf{x} \in \omega_j, j \in \{1, 2, 3\} \quad (2.5)$$

- Moyenne dans le temps de l'écart-type de la population à l'échelle des sous-domaines $\omega_1, \omega_2, \omega_3$ constitués par les 3 types d'occupation du sol

$$\sigma_{\omega_j}(t) = \frac{1}{T} \sum_{t=1}^T \sqrt{\frac{1}{N_j} \sum_{i=1}^{N_j} (P(\mathbf{x}_i, t) - \overline{P(t)})^2} \mathbf{x} \in \omega_j, j \in \{1, 2, 3\} \quad (2.6)$$

2.2.4 Classification des modèles selon leur réponse à l'hétérogénéité du paysage

L'objectif ici est de constituer des groupes de modèles sur la base des similarités entre leurs motifs (le vecteur de descripteurs décrits en 2.2.3). En statistiques, cette tâche relève du partitionnement de données ("*clustering*"), qui consiste à grouper un ensemble d'objets de manière à ce que les individus d'un même groupe ("*cluster*") soient plus "similaires" entre eux qu'avec les individus des autres groupes. Cette tâche repose donc sur la notion de distance (ou dissimilarités) entre les individus, l'objectif étant de minimiser la distance entre les individus d'une même classe et de maximiser la distance entre les individus de classes différentes.

Dans le cadre de notre étude, il est pertinent de déterminer des modèles de processus dont les motifs sont représentatifs de motifs issus d'autres modèles. L'algorithme de classification des *K-medoids* (Kaufmann et Rousseeuw 1987) permet de classer les individus (les modèles de processus) autour de l'individu le plus représentatif de la classe. Chaque classe est donc représentée par un de ses objets (le médoïde), et on peut ainsi comparer plus finement les signatures de population des médoïdes qu'on ne l'aurait fait avec l'ensemble des modèles.

Une difficulté inhérente à tout problème de classification est la définition du nombre de classes optimal. Parmi les méthodes disponibles, la méthode des silhouettes moyennes (Rousseeuw 1987) est la plus classique. La valeur de la silhouette est calculée pour chaque individu, elle quantifie l'adéquation de l'individu à la classe à laquelle il est affecté en comparant la distance de l'individu avec les individus

partageant sa classe et ceux des autres classes. La moyenne des silhouettes est alors calculée pour chaque valeur possible de classes (*i.e.* de 2 à $I - 1$, I étant le nombre d'individus à classer). Le nombre optimal de classes est alors pris comme celui générant la silhouette moyenne la plus élevée.

On définit de cette manière le modèle de classification optimal pour chaque valeur d'hétérogénéité du paysage. Pour définir quels processus écologiques sont les plus proches, on dénombre alors le nombre de fois où les processus deux à deux sont classés dans le même cluster. On définit alors un écotype comme l'ensemble des processus classés dans le même cluster quelle que soit la valeur de la variable paysagère.

2.2.5 Caractérisation des réponses des écotypes à l'hétérogénéité du paysage

À ce stade, les écotypes et les modèles de processus les plus représentatifs de ces écotypes ont été déterminés. On cherche ici à déterminer quelles métriques de population sont responsables des divergences entre les écotypes dans leur réponse à l'hétérogénéité du paysage.

Pour chaque modèle représentatif des différents écotypes (*i.e.* pour chaque médoïde), un modèle linéaire entre les métriques de population et la valeur de densité d'interfaces a été ajusté.

$$\phi_{\mathcal{M}_i} = aED + b \tag{2.7}$$

On peut alors définir chaque médoïde par un vecteur de longueur u , prenant comme valeur $\{-1,1,0\}$ selon que la relation entre la métrique et la densité d'interface est significativement négative, positive ou non significative.

3

Résultats et discussion

3.1 Génération des paysages virtuels

La durée de génération d'une instance de paysage varie entre 3 et 5 minutes sur un processeur de 2 GHz, dépendant de la valeur cible d'ED, les paysages avec des parcelles plus petites et donc plus nombreuses étant plus longues à modéliser.

Du fait de la régularité des distances entre les germes du processus de points sous-jacent, les surfaces des polygones présentent une faible variabilité (Fig. 3.1).

En tolérant un écart de 5% à la valeur cible de $1/3$ de la surface pour les 3 types d'occupation du sol (3.2 (b)), l'équiproportion est respectée sauf pour la plus faible valeur de densités d'interfaces. Cela peut s'expliquer par le fait qu'il est plus difficile de faire des ajustements fins des surfaces lorsque les surfaces des polygones sont élevées.

Concernant les valeurs d'ED, le modèle permet de générer des instances de paysages présentant des valeurs cibles allant de 4 à 25, embrassant ainsi les valeurs rencontrées dans le paysage réel. Du fait de la contrainte sur le nombre de polygones par paysages, il est impossible d'atteindre des valeurs d'ED supérieures à 25 (figure 3.2 (a)). Cela est du au fait que le pavage de Voronoï, additionné à la régularité du processus de points de MSHC contraint la compacité des patches. En somme, il existe une valeur seuil du ratio périmètre/aire qui ne peut être surpassée sans modifier la forme des patches. D'autres méthodes de tessellations, comme les tessellations en T (Kiêu *et al.* 2013) ou les tessellations rectangulaires (Le Ber *et al.* 2009) permettraient de donner des formes moins compactes aux polygones et donc de représenter de manière plus réaliste la forme des parcelles agricoles.

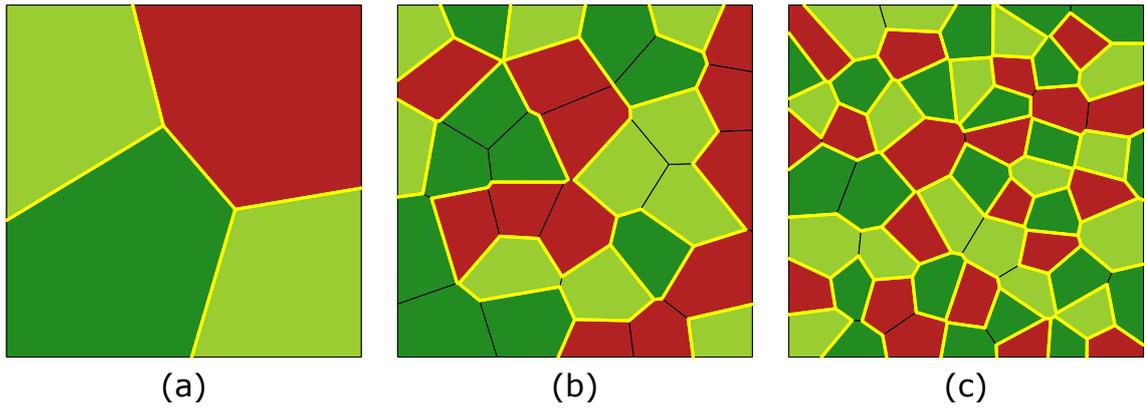


FIGURE 3.1 – Exemples de paysages générés. Les interfaces en jaune (interfaces entre polygones de types différents) sont celles utilisées pour le calcul d’ED. (a) ED=4 , (b) ED=14 et (c) ED=24.

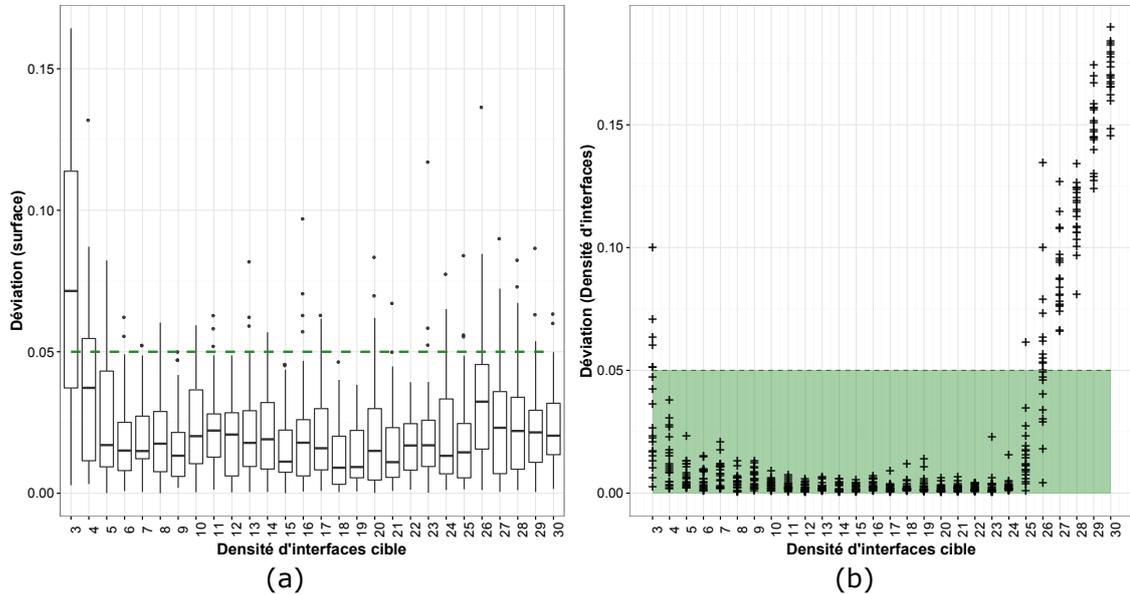


FIGURE 3.2 – Écart entre valeurs cibles et valeurs atteintes pour les paysages générés (a) par rapport à la valeur de $1/3$ de la surface totale du paysage (b) par rapport à la valeur de densité d’interfaces

3.2 Typologie des modèles de processus sur la base de leur signature populationnelle

La classification (le nombre de classes et la manière dont sont regroupés les modèles de processus) diffère selon la valeur d’hétérogénéité du paysage. Les nombres de classes, définis à l’aide de la méthode des silhouettes, sont de 4, 2, 2, 3, et 3 pour les valeurs de densités d’interfaces de 4, 9, 14, 19, et 24, respectivement (fig. A.1 en annexe). Le regroupement des modèles selon leurs composantes de réaction ou de diffusion n’est évident que pour les paysages peu hétérogènes. En effet les quatre classes sont composées selon la composante de démographie : i) Réaction hétérogène

et constante (IC), rassemblés autour du médoïde $IC - HC$, ii) réaction homogène et constante (HC), homogène et décroissante dans le temps (HD) ou nulle ($NULL$), médoïde $HC - IC$, iii) réaction homogène et variant dans le temps avec un pic en milieu de saison (HV), médoïde $HV - IV$, iv) Réaction hétérogène et variant dans le temps (IV), médoïde $IV - HD$. Lorsque le paysage est faiblement hétérogène, la composante de diffusion a donc une influence plus faible que la composante de démographie. Lorsque les interfaces sont moins denses, les patches sont de taille plus importantes et les effets de voisinage sont moins impactants à l'échelle du paysage.

Nous avons ensuite identifié les tuples de modèles de processus qui se trouvent toujours classifiés ensemble quelle que soit la valeur d'hétérogénéité du paysage. Chaque tuple constitue un écotype. En ré-effectuant une classification sur les distances agrégées en contraignant un nombre de classes égal au nombre d'écotypes, nous avons pu identifier les modèles de processus les plus représentatifs des écotypes (médoïdes). La figure 3.3 rassemble les modèles par écotypes et montre les descripteurs des populations associés. Les 6 écotypes sont les suivants :

- type 1 : modèles à réaction homogène ou nulle ($HC, HD, HV, NULL$) et diffusion hétérogène et constante (IC)
- type 2 : réaction homogène ou nulle ($HC, HD, HV, NULL$) et diffusion hétérogène et variant dans le temps (IV); modèle à réaction et diffusion hétérogène et variant dans le temps ($IV-IV$); modèle à réaction hétérogène et variant dans le temps et diffusion nulle ($IV-NULL$)
- type 3 : modèles à réaction hétérogène et constante, diffusion homogène ($IC-HC, IC-HD, IC-HV$); modèle à réaction hétérogène et constante, diffusion hétérogène et variant dans le temps ($IC-IV$)
- type 4 : modèle à réaction et diffusion hétérogène et constante ($IC-IC$); modèle à réaction hétérogène et constante, diffusion nulle ($IC-NULL$)
- type 5 : Modèles à réaction hétérogène et variant dans le temps, diffusion homogène ($IV-HC, IV-HD, IV-HV$)
- type 6 : Modèle à réaction hétérogène et variant dans le temps, diffusion hétérogène et constante ($IV-IC$)

À ce stade, le nombre d'hypothèses en compétition sur les processus écologiques, initialement de 20, a été réduit à 6 classes, chacune représentative d'un type de motif de population, et regroupant en son sein tous les processus présentant des signatures similaires quelle que soit la valeur de la variable paysagère. En jouant sur la structure paysagère dans une optique de favoriser l'organisme d'intérêt, 6 réponses en termes de population sont possibles sachant le jeu d'hypothèses initial sur les processus écologiques.

La connaissance sur les processus écologiques de *P. melanarius* doivent donc être

complétées avant de préconiser des aménagements du paysage en faveur de sa dynamique de population. La démarche à adopter serait d'échantillonner *P. melanarius* sur le terrain et de comparer les motifs obtenus avec les motifs simulés. Le processus simulé qui générera des motifs similaires aux motifs observés pourra être considéré comme le processus le plus à même de représenter les dynamiques spatio-temporelles de *P. melanarius*.

L'inférence des processus écologiques à partir de l'observation des motifs sur le terrain amène cependant de nouveaux questionnements sur l'échantillonnage : comment conduire l'échantillonnage sur le terrain (*i.e.* où et quand échantillonner, quel volume de données est nécessaire) pour maximiser la robustesse de l'inférence des processus ? Ces questions sur l'échantillonnage dans une optique d'inférence des processus sont l'objet du chapitre III.

3.3 Relation entre les métriques de population et la variable paysagère

La figure 3.4 montre la relation entre les valeurs de métriques de population agrégées en temps des modèles médoïdes des 6 écotypes. Les variables sd_1 , sd_2 et sd_3 , moyenne en temps de l'écart-type dans les habitats de qualité "moyenne", "bonne" et "mauvaise", respectivement, ne sont pas représentées car leur relation linéaire avec la variable de paysage est non significative, quel que soit le modèle de processus considéré. Pour chacune des métriques représentées, la variable paysagère explique au moins 70% de la variance ($R^2 \geq 0.7$).

On remarque que quel que soit le processus considéré, la densité d'interface a une influence négative ou nulle sur la population moyenne à l'échelle du paysage. Pour les autres critères, la direction de la réponse à l'hétérogénéité du paysage dépend du processus considéré.

Selon l'unique critère de moyenne de population dans le temps à l'échelle du paysage, augmenter l'hétérogénéité de configuration (augmenter la densité d'interfaces, faire de petites parcelles,...) a donc un effet négatif. La préconisation en termes d'aménagement du paysage, au vu du set initial d'hypothèses sur les processus écologiques de l'organisme d'intérêt, serait de conserver des parcelles de grande taille (relativement à la taille des parcelles de la ZAA, de l'ordre de 3 ha) avec une densité d'interfaces 133 m/ha. Ces conclusions sont en contradiction avec l'idée communément admise que l'hétérogénéité de configuration a un effet bénéfique sur les dynamiques de population (Fahrig, Girard *et al.* 2015). Nous discuterons cet aspect plus en détail dans le chapitre 4.

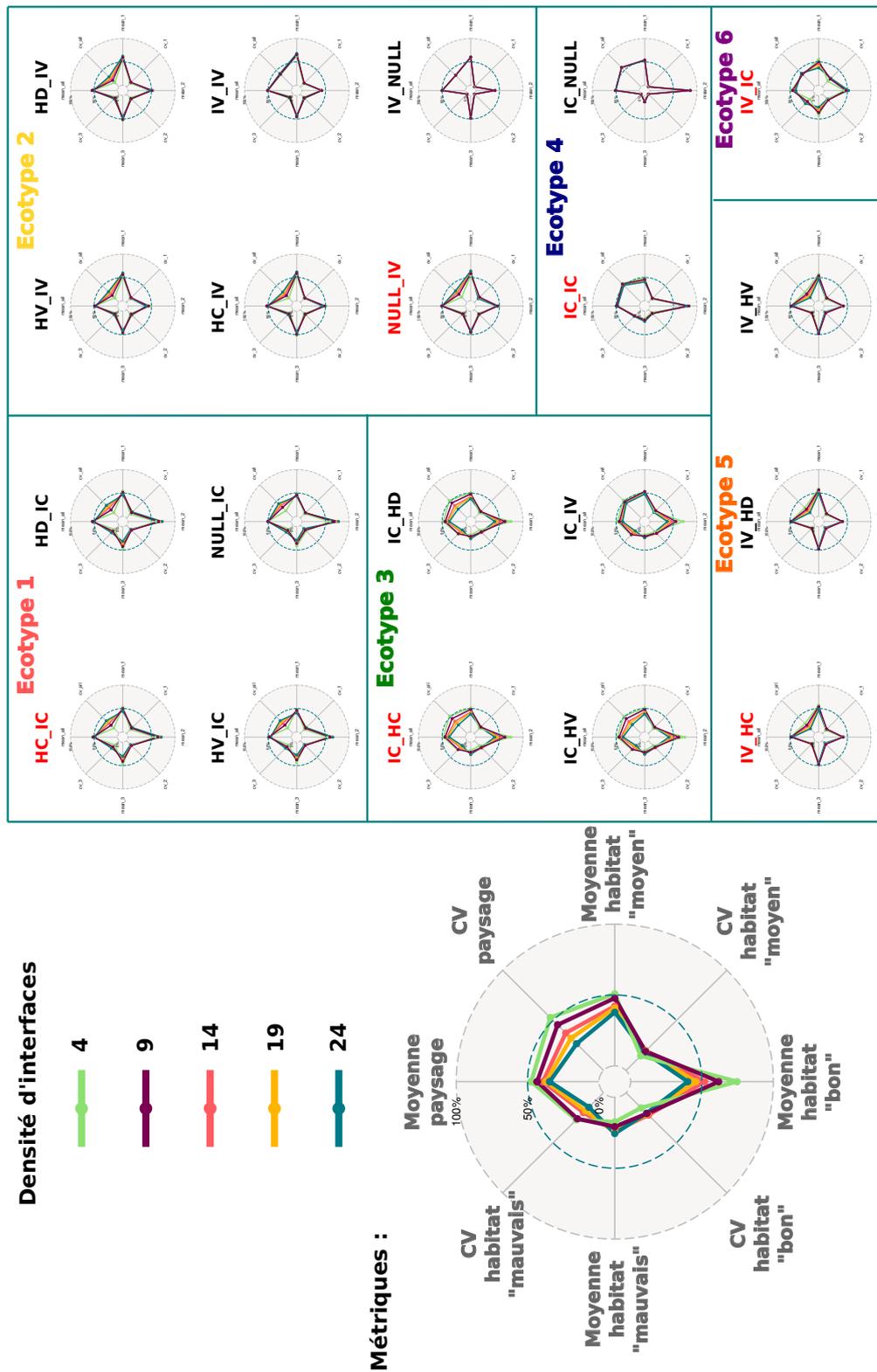


FIGURE 3.3 – Regroupement des modèles de processus en "écotypes". Des modèles appartenant aux mêmes écotypes sont classifiés dans le même cluster quelle que soit la valeur d'hétérogénéité du paysage. Les figures "radar" représentent la valeur des métriques moyennées dans le temps par valeur d'hétérogénéité du paysage. CV = coefficient de variation.

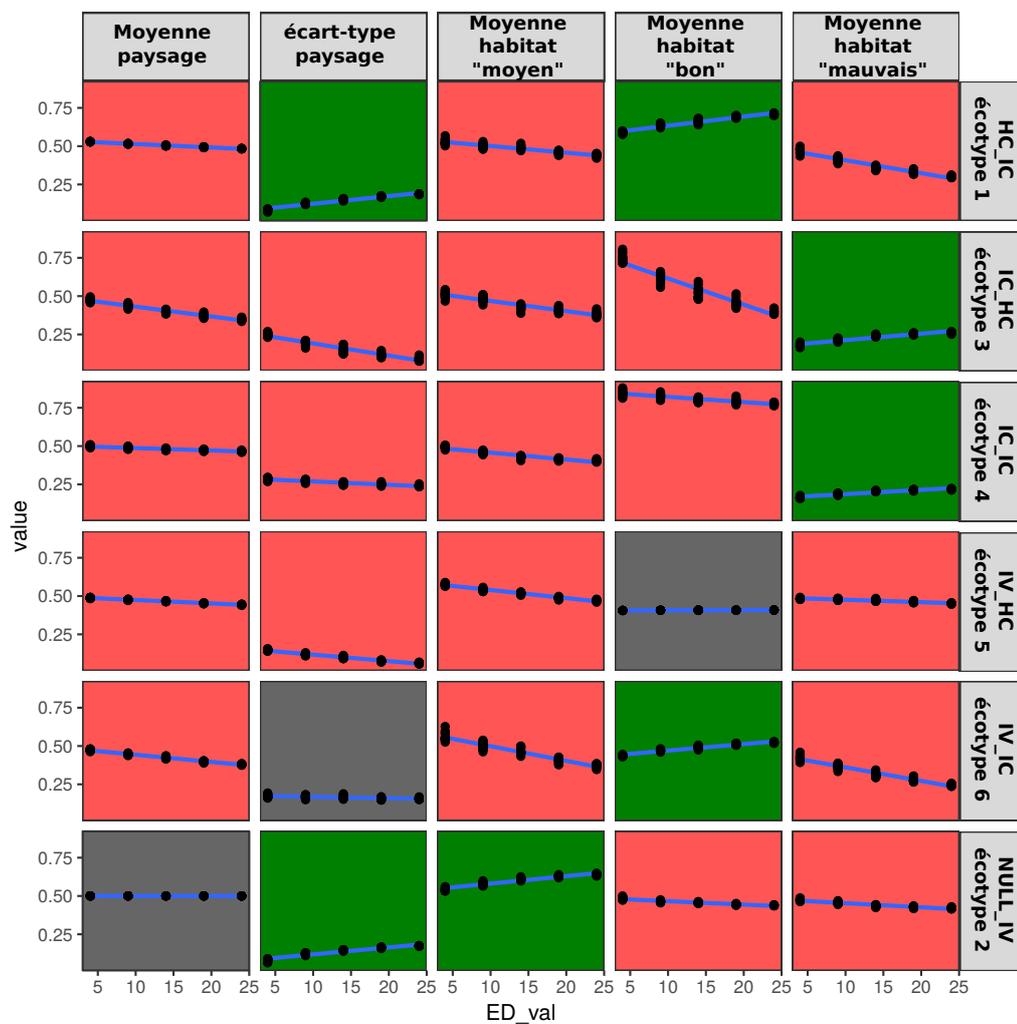


FIGURE 3.4 – Représentation des valeurs de métriques de population le long du gradient de la variable paysagère pour chacun des modèles médoïdes. La couleur de l'arrière-plan (rouge, vert, gris) caractérise une relation significativement négative, positive ou non significative avec la densité d'interfaces.

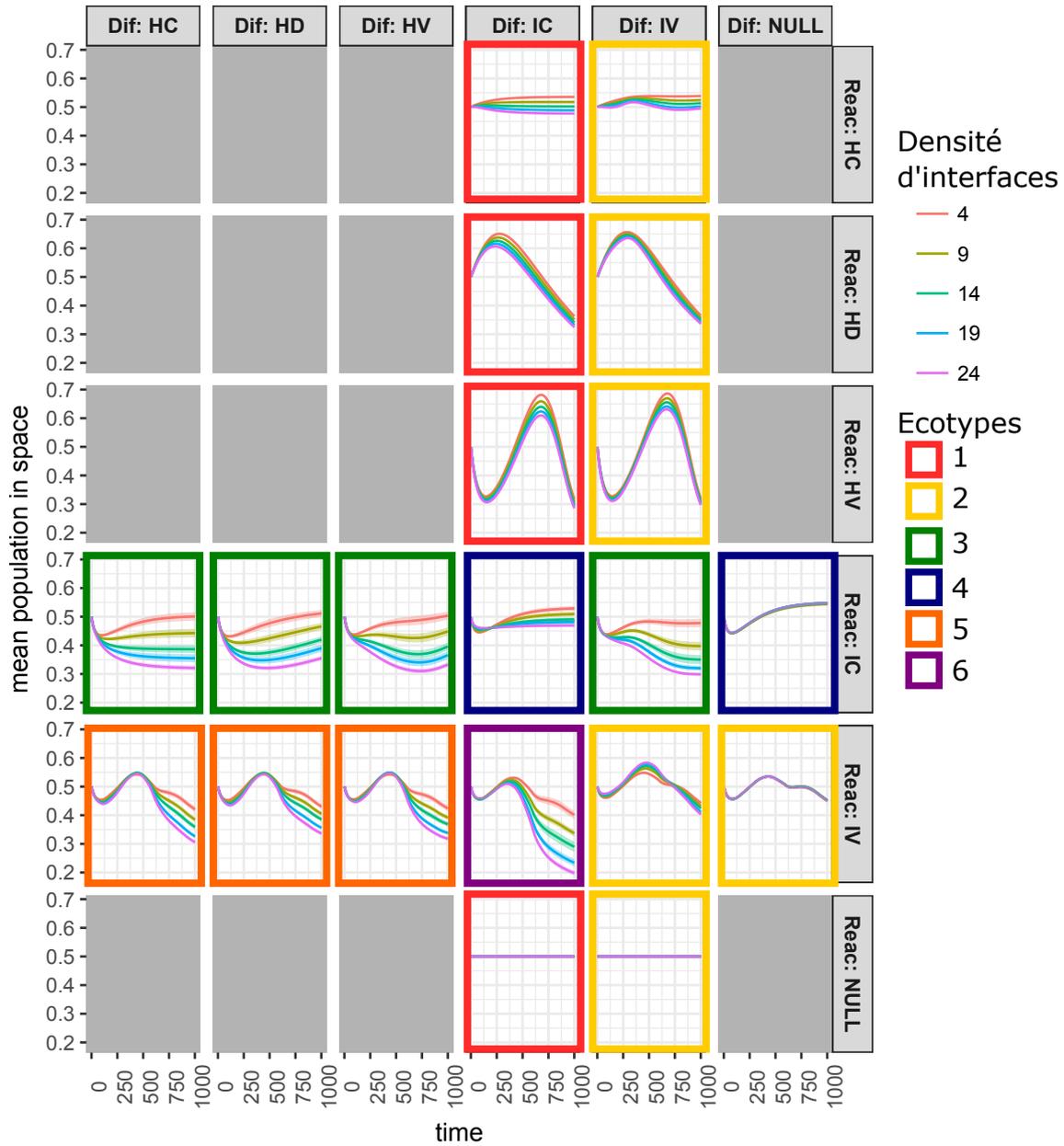


FIGURE 3.5 – Suivi temporel de la population moyenne à l'échelle du paysage pour tous les modèles. Les couleurs des encadrés correspondent au clusters formés sur la base des motifs de population.

4

Conclusion du chapitre 2

Dans ce chapitre, nous avons mis en place une heuristique permettant la classification des processus écologiques sur la base des motifs de population qu'ils génèrent. Dans un premier temps, nous avons créé un modèle de paysage permettant de générer des instances de domaines spatiaux présentant des propriétés recherchées en termes de longueurs d'interfaces et de proportion des différents types de couverts. Nous avons ensuite formalisé et simulé sur ces paysages virtuels des modèles représentatifs d'hypothèses en compétition concernant les processus écologiques régissant les dynamiques de population d'un auxiliaire de culture. Ces modèles de processus ont été classifiés sur la base de leurs motifs pour chaque valeur de la variable paysagère d'intérêt.

En comparant les classifications associées à chaque valeur de la variable paysagère, nous avons pu déterminer des *écotypes* qui sont des classes de modèles de processus qui possèdent chacune une réponse spécifique à la variable d'hétérogénéité de configuration. La réduction du nombre d'hypothèses initiales en un nombre restreint d'hypothèses permet d'avancer dans la démarche de l'amélioration de la connaissance des relations processus-motifs qui permettra, à terme, de définir des scénarios d'aménagement favorables à des espèces ou des peuplements d'intérêts. Le chapitre III est l'étape suivante dans ce processus d'acquisition de connaissances et montre comment planifier au mieux l'échantillonnage sur le terrain pour favoriser la sélection du processus parmi cet ensemble irréductible d'hypothèses. Si l'heuristique définie dans ce chapitre l'a été à partir des connaissances écologiques de *P. melanarius*, son usage est généralisable à toute espèce pour lesquelles des hypothèses sur les processus écologiques sont formulables *a priori*.

Références de la partie II

- Alema, Bas (2014). “Quantifying and simulating movement of the predator carabid beetle *Pterostichus melanarius* in arable land”. Thèse de doct. Wageningen : Wageningen University. URL : <https://www.wageningenur.nl/en/Publication-details.htm?publicationId=publication-way-343533353438> (cf. pp. 25, 31, 33, 34, 64, 66, 82, 101).
- Baars, M. A. (12/1979). “Patterns of movement of radioactive carabid beetles”. en. In : *Oecologia* 44.1, p. 125-140. URL : <https://link.springer.com/article/10.1007/BF00346411> (cf. p. 24).
- Baddeley, Adrian (12/2010). *Analysing spatial point patterns in 'R'*. Australia. URL : <http://www.csiro.au/files/files/pn0y.pdf> (cf. p. 30).
- Banks, Robert B. (1994). *Growth and Diffusion Phenomena*. Sous la dir. de F. John, J. E. Marsden, L. Sirovich, M. Golubitsky et W. Jäger. T. 14. Texts in Applied Mathematics. Berlin, Heidelberg : Springer Berlin Heidelberg. URL : <http://link.springer.com/10.1007/978-3-662-03052-3> (cf. pp. 34, 65).
- Barros, Ceres, Stephen C. F. Palmer, Greta Bocedi et Justin M. J. Travis (12/2016). “Spread rates on fragmented landscapes : the interacting roles of demography, dispersal and habitat availability”. en. In : *Diversity and Distributions* 22.12, p. 1266-1275. URL : <http://onlinelibrary.wiley.com/doi/10.1111/ddi.12487/abstract> (cf. pp. 34, 65).
- Bommarco, Riccardo, David Kleijn et Simon G. Potts (04/2013). “Ecological intensification : harnessing ecosystem services for food security”. In : *Trends in Ecology & Evolution* 28.4, p. 230-238. URL : <http://www.sciencedirect.com/science/article/pii/S016953471200273X> (cf. pp. 20, 24, 103).
- Bourhis, Yoann, Sylvain Poggi, Youcef Mammeri, Anne-Marie Cortesero, Anne Le Ralec et Nicolas Parisey (09/2015). “Perception-based foraging for competing resources : Assessing pest population dynamics at the landscape scale from heterogeneous resource distribution”. In : *Ecological Modelling* 312, p. 211-221. URL : <http://www.sciencedirect.com/science/article/pii/S0304380015002355> (cf. pp. 34, 65, 121).
- CÁrcamo, Hector A. et John R. Spence (06/1994). “Crop Type Effects on the Activity and Distribution of Ground Beetles (Coleoptera : Carabidae)”. In : *Environmental Entomology* 23.3, p. 684-692. URL : <https://academic.oup.com/ee/article/23/3/684/417992> (cf. p. 24).
- Cushman, Samuel A., Kevin McGarigal et Maile C. Neel (09/2008). “Parsimony in landscape metrics : Strength, universality, and consistency”. In : *Ecological Indicators* 8.5, p. 691-703. URL : <http://www.sciencedirect.com/science/article/pii/S1470160X07001306> (cf. pp. 26, 28, 64).
- Duelli, Peter (04/1997). “Biodiversity evaluation in agricultural landscapes : An approach at two different scales”. In : *Agriculture, Ecosystems & Environment*. Biodiversity in Agriculture - for a Sustainable Future 62.2, p. 81-91. URL : <http://www.sciencedirect.com/science/article/pii/S0167880996011437> (cf. pp. 2, 25).
- Edelman, Andrew J. (02/2012). “Positive Interactions between Desert Granivores : Localized Facilitation of Harvester Ants by Kangaroo Rats”. In : *PLoS ONE* 7.2. URL : <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3279350/> (cf. pp. 29, 30, 65).

- Fahrig, Lenore, Jacques Baudry, Lluís Brotons, Françoise G. Burel, Thomas O. Crist, Robert J. Fuller, Clelia Sirami, Gavin M. Siriwardena et Jean-Louis Martin (02/2011). “Functional landscape heterogeneity and animal biodiversity in agricultural landscapes”. en. In : *Ecology Letters* 14.2, p. 101-112. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2010.01559.x/abstract> (cf. pp. 2, 26, 59, 64).
- Fahrig, Lenore, Judith Girard, Dennis Duro, Jon Pasher, Adam Smith, Steve Javorek, Douglas King, Kathryn Freemark Lindsay, Scott Mitchell et Lutz Tischendorf (02/2015). “Farmlands with smaller crop fields have higher within-field biodiversity”. In : *Agriculture, Ecosystems & Environment* 200, p. 219-234. URL : <http://www.sciencedirect.com/science/article/pii/S0167880914005192> (cf. pp. 7, 9, 26, 43, 59, 63).
- Fahrig, Lenore et William K. Nuttle (2005). “Population Ecology in Spatially Heterogeneous Environments”. en. In : *Ecosystem Function in Heterogeneous Landscapes*. Sous la dir. de Gary M. Lovett, Monica G. Turner, Clive G. Jones et Kathleen C. Weathers. DOI : 10.1007/0-387-24091-8_6. Springer New York, p. 95-118. URL : http://link.springer.com/chapter/10.1007/0-387-24091-8_6 (cf. pp. 25, 26).
- Gaucherel, C. (11/2008). “Neutral models for polygonal landscapes with linear networks”. In : *Ecological Modelling* 219.1–2, p. 39-48. URL : <http://www.sciencedirect.com/science/article/pii/S0304380008003803> (cf. pp. 28, 64).
- Geiger, Flavia *et al.* (03/2010). “Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland”. In : *Basic and Applied Ecology* 11.2, p. 97-105. URL : <http://www.sciencedirect.com/science/article/pii/S1439179109001388> (cf. p. 24).
- Hallmann, Caspar A. *et al.* (2017). “More than 75 percent decline over 27 years in total flying insect biomass in protected areas”. eng. In : *PloS One* 12.10, e0185809 (cf. p. 20).
- Holland J. M, Begbie M. (2004). “The spatial dynamics and movement of *Pterostichus melanarius* and *P. madidus* (Carabidae) between and within arable fields in the UK”. In : *International Journal of Ecology and Environmental Sciences* 30, p. 35-50 (cf. pp. 24, 61).
- Holland, J. M., C. F. G. Thomas, T. Birkett, S. Southway et H. Oaten (12/2005). “Farm-scale spatiotemporal dynamics of predatory beetles in arable crops”. en. In : *Journal of Applied Ecology* 42.6, p. 1140-1152. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2664.2005.01083.x/abstract> (cf. pp. 21, 64, 67).
- Holling, C. S. et Craig R. Allen (06/2002). “Adaptive Inference for Distinguishing Credible from Incredible Patterns in Nature”. en. In : *Ecosystems* 5.4, p. 319-328. URL : <https://link.springer.com/article/10.1007/s10021-001-0076-2> (cf. pp. 9, 35, 96).
- Honek, A. (1988). “The effect of crop density and microclimate on pitfall trap catches of Carabidae, Staphylinidae (Coleoptera), and Lycosidae (Araneae) in cereal fields”. In : *Pedobiologia (Jena)* 32.3-4, p. 233-242. URL : <http://www.refdoc.fr/Detailnotice?idarticle> (cf. pp. 24, 101).
- Illian, Dr Janine, Prof Antti Penttinen, Dr Helga Stoyan et Dr Dietrich Stoyan (04/2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. en. John Wiley & Sons (cf. pp. 28-30).

- Johnson, Jerald B. et Kristian S. Omland (02/2004). “Model selection in ecology and evolution”. In : *Trends in Ecology & Evolution* 19.2, p. 101-108. URL : <http://www.sciencedirect.com/science/article/pii/S0169534703003458> (cf. pp. 35, 60, 69, 83).
- Kaufmann, Leonard et Peter Rousseeuw (01/1987). “Clustering by Means of Medoids”. In : *Data Analysis based on the L1-Norm and Related Methods*, p. 405-416 (cf. p. 38).
- Kiêu, Kiên, Katarzyna Adamczyk-Chauvat, Hervé Monod et Radu S. Stoica (02/2013). “A completely random T-tessellation model and Gibbsian extensions”. In : *arXiv :1302.1809 [math, stat]*. arXiv : 1302.1809. URL : <http://arxiv.org/abs/1302.1809> (cf. p. 40).
- Kromp, Bernhard (06/1999). “Carabid beetles in sustainable agriculture : a review on pest control efficacy, cultivation impacts and enhancement”. In : *Agriculture, Ecosystems & Environment* 74.1-3, p. 187-228. URL : <http://www.sciencedirect.com/science/article/pii/S0167880999000377> (cf. pp. 24, 100).
- Landis, Douglas A., Stephen D. Wratten et Geoff M. Gurr (2000). “Habitat Management to Conserve Natural Enemies of Arthropod Pests in Agriculture”. In : *Annual Review of Entomology* 45.1, p. 175-201. URL : <http://dx.doi.org/10.1146/annurev.ento.45.1.175> (cf. pp. 5, 24, 59).
- Lausch, Angela, Thomas Blaschke, Dagmar Haase, Felix Herzog, Ralf-Uwe Syrbe, Lutz Tischendorf et Ulrich Walz (01/2015). “Understanding and quantifying landscape structure – A review on relevant process characteristics, data models and landscape metrics”. In : *Ecological Modelling. Use of ecological indicators in models* 295, p. 31-41. URL : <http://www.sciencedirect.com/science/article/pii/S0304380014003974> (cf. p. 28).
- Le Ber, Florence, Claire Lavigne, Katarzyna Adamczyk, Frédérique Angevin, Nathalie Colbach, Jean-François Mari et Hervé Monod (2009). “Neutral modelling of agricultural landscapes by tessellation methods-Application for gene flow simulation.” In : *Ecological Modelling* 220, p. 3536-3545. URL : <https://hal.archives-ouvertes.fr/hal-00409081> (cf. pp. 28, 40, 64).
- Marrec, Ronan, Isabelle Badenhauer, Vincent Bretagnolle, Luca Börger, Marilyn Roncoroni, Nadine Guillon et Bertrand Gauffre (01/2015). “Crop succession and habitat preferences drive the distribution and abundance of carabid beetles in an agricultural landscape”. In : *Agriculture, Ecosystems & Environment* 199, p. 282-289. URL : <http://www.sciencedirect.com/science/article/pii/S016788091400468X> (cf. pp. 7, 9, 24, 33, 67).
- McGarigal, Kevin (2006). “Landscape Pattern Metrics”. en. In : *Encyclopedia of Environmetrics*. John Wiley & Sons, Ltd. URL : <http://onlinelibrary.wiley.com/doi/10.1002/9780470057339.val006.pub2/abstract> (cf. p. 28).
- Padrón, Víctor et María Cristina Trevisan (02/2006). “Environmentally induced dispersal under heterogeneous logistic growth”. In : *Mathematical Biosciences* 199.2, p. 160-174. URL : <http://www.sciencedirect.com/science/article/pii/S0025556405002075> (cf. pp. 34, 65).
- Parisey, Nicolas, Yoann Bourhis, Lionel Roques, Samuel Soubeyrand, Benoît Ricci et Sylvain Poggi (12/2016). “Rearranging agricultural landscapes towards habitat quality optimisation : In silico application to pest regulation”. In : *Ecological Complexity* 28, p. 113-122. URL : <https://www.sciencedirect.com/science/article/pii/S1476945X16300538> (cf. pp. 5, 34, 59, 101).
- Press, William H. (2007). *Numerical recipes 3rd edition : The art of scientific computing*. Cambridge university press. URL : <https://books.google.fr/books?hl=fr&lr=&id=1aA0dzK3FegC&oi=>

[fnd&pg=PA1&dq=press+numerical+recipes&ots=3jRmG8Grog&sig=C3unnUCUA2krwf0mzPVkAslLcKY](#)
(cf. p. 37).

- Rijnsdorp, A. D. (05/1980). "Pattern of movement in and dispersal from a dutch forest of *Carabus problematicus* Hbst. (Coleoptera, Carabidae)". eng. In : *Oecologia* 45.2, p. 274-281 (cf. p. 24).
- Roques, Lionel, Marie-Anne Auger-Rozenberg et Alain Roques (11/2008). "Modelling the impact of an invasive insect via reaction-diffusion". In : *Mathematical Biosciences* 216.1, p. 47-55. URL : <http://www.sciencedirect.com/science/article/pii/S0025556408001235> (cf. pp. 33, 65).
- Roques, Lionel et Samuel Soubeyrand (2013). *Modèles de réaction-diffusion pour l'écologie spatiale*. français. Versailles, France : Ed. Quae, impr. 2013 (cf. pp. 33, 34).
- Rousseeuw, Peter J. (11/1987). "Silhouettes : A graphical aid to the interpretation and validation of cluster analysis". In : *Journal of Computational and Applied Mathematics* 20.Supplement C, p. 53-65. URL : <http://www.sciencedirect.com/science/article/pii/0377042787901257> (cf. p. 38).
- Shigesada, Nanako et Kohkichi Kawasaki (02/1997). *Biological Invasions : Theory and Practice*. en. Oxford University Press, UK (cf. pp. 33, 65).
- Sunderland, K.D. (01/2002). "Invertebrate pest control by carabids". In : *The Agroecology of Carabid Beetles*, p. 165-214 (cf. p. 24).
- Thiele, Hans Ulrich (1977). *Carabid beetles in their environments : a study on habitat selection by adaptations in physiology and behaviour*. en. Google-Books-ID : bK_wAAAAMAAJ. Springer-Verlag (cf. p. 24).
- Thomas, C. F. G., N. J. Brown et D. A. Kendall (04/2006). "Carabid movement and vegetation density : Implications for interpreting pitfall trap data from split-field trials". In : *Agriculture, Ecosystems & Environment* 113.1-4, p. 51-61. URL : <http://www.sciencedirect.com/science/article/pii/S0167880905004330> (cf. p. 24).
- Vasseur, Chloé, Alexandre Joannon, Stéphanie Aviron, Françoise Burel, Jean-Marc Meynard et Jacques Baudry (02/2013). "The cropping systems mosaic : How does the hidden heterogeneity of agricultural landscapes drive arthropod populations?" In : *Agriculture, Ecosystems & Environment*. Landscape ecology and biodiversity in agricultural landscapes 166, p. 3-14. URL : <http://www.sciencedirect.com/science/article/pii/S0167880912003246> (cf. pp. 2, 25).
- Wallin, H. et B. S. Ekbom (10/1988). "Movements of carabid beetles (Coleoptera : Carabidae) inhabiting cereal fields : a field tracing study". en. In : *Oecologia* 77.1, p. 39-43. URL : <http://link.springer.com/article/10.1007/BF00380922> (cf. p. 24).
- Wallin, Henrik et Barbara Ekbom (10/1994). "Influence of Hunger Level and Prey Densities on Movement Patterns in Three Species of Pterostichus Beetles (Coleoptera : Carabidae)". en. In : *Environmental Entomology* 23.5, p. 1171-1181. URL : <http://ee.oxfordjournals.org/content/23/5/1171> (cf. p. 24).
- Winder, L., G. J. K. Griffiths, J. N. Perry, C. J. Alexander, J. M. Holland, P. J. Kennedy et A. Birt (12/2005). "The role of large-scale spatially explicit and small-scale localized processes on

the population dynamics of cereal aphids". eng. In : *Bulletin of Entomological Research* 95.6, p. 579-587 (*cf.* p. 24).

Troisième partie

Inférence des processus écologiques par l'analyse des motifs de population : un guide méthodologique pour l'échantillonnage dans l'espace et le temps

1	Préface à l'article	54
1.1	Présentation générale de la démarche	54
1.2	Remarque sur les hypothèses écologiques sélectionnées	56
2	Article soumis à <i>Ecological Modelling</i>	57
2.1	Introduction	59
2.2	Material and Methods	61
2.3	Results	70
2.4	Discussion and conclusions	81
3	Perspectives d'amélioration concernant le mode d'échantillonnage	85
	Références	94

1

Préface à l'article

1.1 Présentation générale de la démarche

Les travaux que nous présentons dans ce chapitre se positionnent dans le champ thématique de l'inférence des mécanismes écologiques (*processes*) à partir de l'étude des motifs (*patterns*) de population dans l'espace et le temps. Ces travaux ont été soumis en tant qu'article en novembre 2017, accepté sous réserve de modifications majeures en décembre 2017. L'idée est que les processus écologiques amènent à des *patterns* spatio-temporels de population spécifiques de ces processus (McIntire et Fajardo 2009). En étudiant les *patterns*, il est théoriquement possible d'inférer les processus qui les ont générés, et de nombreuses études sont parvenues à inférer les traits de vie d'une espèce à partir de l'étude de *patterns* de population, en particulier dans le domaine de l'écologie du mouvement (Nathan 2008 ; McIntire, Rompré *et al.* 2013). Cependant, comme nous en avertissent McIntire et Fajardo (2009), l'inférence des mécanismes peut être erronée si la description des motifs de population est insuffisante, *i.e.* si les données collectées pour décrire la population sont trop peu nombreuses ou de mauvaise qualité. Ce constat pose la question de l'échantillonnage des données biologiques dans l'espace et le temps : quel est l'échantillonnage nécessaire à mener sur le terrain pour que l'inférence des processus écologiques à partir des motifs de population soit possible ?

Nos travaux posent une heuristique pour aborder cette question. L'approche utilisée est celle dite de "l'écologue virtuel" (Zurell *et al.* 2010), qui génère des jeux de données à partir de simulations spatialement explicites de modèles de dynamiques de population, et définit des stratégies d'échantillonnage sur la base de ces données d'observations virtuelles. L'objectif de ces travaux est de définir *a priori* des stratégies d'échantillonnage qui permettent de maximiser la robustesse de l'inférence des processus écologiques sur la base des motifs observés. La démarche méthodologique générale, expliquée plus en détail dans l'article, est la suivante :

1. **Formulation d'hypothèses écologiques en compétition concernant les processus écologiques et traduction de ces hypothèses en modèles de dynamiques de population.**

Pour illustrer nos travaux, nous avons formulé des hypothèses alternatives plau-

sibles sur les relations entre l'hétérogénéité de deux traits de vie (taux de croissance et capacité de dispersion) et les couverts végétaux constituant le paysage. La multiplicité des hypothèses écologiques traduit notre incertitude sur les mécanismes régissant les dynamiques de population d'un auxiliaire de culture, *Pterostichus melanarius*. Nous avons traduit ces hypothèses en modèles de réaction diffusion.

2. Simulation spatialement explicite des modèles de dynamique des population.

Les modèles de réaction-diffusion ont été simulés sur des paysages virtuels dont les propriétés structurales reflètent celles d'un paysage agricole dominé par la polyculture-élevage en Bretagne, générés à l'aide du modèle de paysage présenté en 3.1.

3. Extraction des motifs de population selon différentes stratégies d'échantillonnages. Les populations virtuelles associées à chaque modèle de processus sont échantillonnées selon différentes modalités, et les motifs spatiaux et/ou temporels sont caractérisés à partir de ces jeux de données.

Nous avons défini des stratégies d'échantillonnage à partir de la littérature et des contraintes d'échantillonnage liées au paysage réel, différant dans les nombres de points d'échantillonnages dans l'espace, de dates d'échantillonnage et de paysages échantillonnés. Nous avons utilisé ces stratégies d'échantillonnages pour échantillonner les populations virtuelles issues des différents modèles, dont nous avons extrait les motifs, ici des statistiques descriptives des dynamiques de population agrégées dans l'espace et le temps.

4. Classification des modèles à partir des motifs de population. Pour chaque stratégie d'échantillonnage, un algorithme de classification est utilisé pour classifier les modèles sur la base des données biologiques échantillonnées. Les performances de classification, *i.e.* la capacité du modèle de classification à correctement classifier les modèles de processus, sont calculées pour chaque stratégie d'échantillonnage. Une bonne performance de classification signifie que les motifs de population associés aux différents modèles de processus sont contrastés, puisqu'ils sont correctement inférés par le modèle de classification. Plus les performances de classification sont élevées, plus la stratégie d'échantillonnage permet de capturer la spécificité de la relation entre les processus écologiques et les motifs de population.

Par ailleurs, décomposer la relation entre les performances de classifications et les différentes dimensions des stratégies d'échantillonnages (nombre de dates, de lieux et de répliqués de paysages échantillonnés) permet de hiérarchiser ces dimensions sur la capacité d'inférence des processus écologiques sur la base

des données échantillonnées.

L'originalité de ces travaux réside avant tout dans la combinaison de l'approche de "l'écologue virtuel" et de l'approche hypothético-inductive (multiplicité des hypothèses sur les processus ; McIntire et Fajardo (2009), Grimm *et al.* (2005) et Platt (1964)) pour définir *a priori* des stratégies d'échantillonnage qui valorisent au maximum les données échantillonnées. Dans le domaine de l'écologie du paysage, où les échelles temporelles et spatiales rendent la collecte de données sur le terrain coûteuse en temps et en ressources humaines et financières, il est crucial de s'assurer que les données échantillonnées permettront une réelle plus-value en termes de connaissances écologiques.

1.2 Remarque sur les hypothèses écologiques sélectionnées

Il convient de préciser que les modèles de dynamique de population présentés dans l'article et les modèles définis comme étant les plus représentatifs des écotypes définis dans le chapitre 2 ne sont pas les mêmes. Trois raisons expliquent cette différence :

- Pour simplifier le message dans le cadre de la publication, les modèles à réaction ou diffusion hétérogène en espace et variant dans le temps (*IV*) ont été omis. Cela avait été intégré dans une première version du manuscrit mais alourdissait considérablement le message puisqu'il imposait de décrire la calibration de ces modèles. Le message étant avant tout méthodologique, intégrer ces hypothèses mécanistes n'avait pas un intérêt suffisant au regard du coût engendré dans la structure de l'article.
- Même si la structure de la thèse ne les présente pas de cette manière, les derniers résultats du chapitre 2 ont été générés chronologiquement après les travaux liés à l'article soumis à *Ecological modelling*. Ce qui a dirigé le choix des hypothèses mécanistes était le fait d'explorer l'ensemble des mécanismes possibles en omettant la modalité *IV*.
- L'idée de représenter des hypothèses écologiques générant des signatures populationnelles proches permet justement d'étudier si la manière dont est conduit l'échantillonnage peut permettre de distinguer ces signatures. Si une stratégie d'échantillonnage permet de discriminer des hypothèses générant des signatures proches, cela montre que l'échantillonnage permet de capturer finement les différences entre les signatures.

Article soumis à *Ecological Modelling*

Infering species ecological traits from population patterns in
space and time –

A simulation-based framework for sampling strategy
selection

Benoit Bellot^{a,b}, Sylvain Poggi^a, Jacques Baudry^b, Yoann Bourhis^a, Nicolas Parisey^a

^aINRA, UMR 1349 IGEPP, Le Rheu, France

^bINRA UMR BAGAP, Rennes, France

Résumé

Une bonne connaissance de la variabilité des traits biologiques d'une espèce en relation à son environnement est la pierre angulaire des études d'aménagement du paysage dans une optique de gestion d'espèce. Une manière d'inférer cette relation est de comparer les motifs de population observés sur le terrain avec des données virtuelles obtenues par simulation spatialement explicite de modèles de dynamiques de population. Cependant la robustesse de l'inférence est tributaire des données d'observation disponible, et donc de la stratégie d'échantillonnage utilisée pour la collecte de ces données. L'échantillonnage sur le terrain est contraint par différents facteurs, comme le nombre de réplicats de paysages, le nombre de visites possibles sur le terrain, et le nombre de pièges, qui doivent être prises en compte avant la campagne d'échantillonnage. Nous présentons et illustrons une méthode heuristique pour aborder la question de l'échantillonnage optimal, conditionné par ces contraintes induites par le paysage.

Nous avons tout d'abord étudié un paysage agricole réel pour déterminer ses propriétés en termes de configuration et de composition, ainsi que le nombre de réplicats de paysages disponibles pour l'échantillonnage. Les propriétés du paysage réel ont été utilisées comme contraintes dans un modèle de paysage pour générer un ensemble de paysages avec des propriétés similaires. Par ailleurs, nous avons formulé des modèles de dynamique de population (notés par la suite *Modèles de Processus PM*) porteurs d'hypothèses en compétition concernant deux processus écologiques

(accroissement de la population et dispersion) en relation avec les variables spatiales pour *Pterostichus melanarius*, une espèce de carabe impliquée dans la régulation des ravageurs. Nous avons simulé ces modèles spatialement explicites et extrait leurs motifs pour chaque stratégie d'échantillonnage. Nous avons défini la qualité des stratégies d'échantillonnage comme sa capacité à capturer les contrastes entre les motifs associés aux différents PM, résumé par les performances d'une procédure de classification. La meilleure stratégie d'échantillonnage a été sélectionnée sur la base des performances de classification et sa faisabilité *in situ*. Nous avons finalement exploré l'effet de l'incertitude liée aux hypothèses écologiques sur les performances de classification par une analyse de sensibilité des paramètres des PM.

S'il reste des améliorations à effectuer avant d'être pleinement opérationnel pour les écologistes du paysage, notre cadre méthodologique contribue à rapprocher la planification expérimentale de son application sur le terrain dans une optique d'inférence des processus écologiques.

Abstract

A good knowledge about species traits variability in relation to their environment is the cornerstone of landscape-oriented species management studies. One way to infer this relationship is to compare species signatures in space and time from field data with spatially explicit population dynamics models outputs. However the inference robustness relies on the available field data, and thus on the quality of the underlying sampling strategy. Field sampling is constrained by several factors, such as the number of landscape replicates, possible number of temporal sessions and number of sample locations, that need to be accounted for prior to field sampling. We set and illustrate a heuristic method to answer the question of optimal sampling conditioned by these landscape-induced constraints. First we studied a real agricultural landscape to determine its properties in terms of configuration and composition, as well as the maximum number of landscape replicates available for the sampling designs set up. The real landscape properties were used as constraints in a landscape model to generate a collection of landscapes with similar properties.

On the other hand, we formulated population dynamics models (hereafter noted *Process models* PM) carrying competing hypotheses about two ecological processes—population growth and dispersal—in relation to spatial covariates for *Pterostichus melanarius*, a carabid species involved in pest regulation. We simulated these spatially explicit models and extracted their sampling-dependent signatures, *i.e.* metrics computed on different population samples. We defined a sampling design quality as its ability to capture the contrasts between the PM signatures, summarised by the performance of a classification procedure. The most relevant sampling design was selected on the basis of classification performance and *in situ* feasi-

lity. Finally we explored the effects of the a priori ecological hypotheses quality on classification performances, through a sensitivity analysis of the PM parameters. While some improvements remain to be achieved before being fully operational for landscape ecologists, our framework contributes to bringing closer sampling theory and its application on the field. It endorses the use of landscape modelling to design sampling prior to field experiment to bring out the best from sampled data.

Keywords

Ecological traits, Spatially Explicit Process Models, Population Dynamics signature, Sampling Designs, Reaction-Diffusion, Classification Algorithm, landscape ecology

2.1 Introduction

One major ambition of applied ecology is to arrange the landscape structure for a purpose of population dynamics management in space and time (Turner 2005). This approach serves a wide variety of purposes, such as conservation of endangered species and biodiversity management (Moloney et Levin 1996; Aben et al. 2016), enhancement of ecosystem services (Landis et al. 2000; Jonsson et al. 2014; Parisey et al. 2016) and control of invasive species (Jongejans et al. 2008; Savage et Renton 2014). An increasingly used approach to tackle these thematics is the exploration of the effects of landscape scenarios on the population dynamics through landscape modelling (Schröder et Seppelt 2006). In this approach, a spatially explicit population model is simulated on landscapes presenting a gradient of one or several properties to be explored (*e.g.* configurational or compositional heterogeneity, *sensu* Fahrig, Baudry et al. (2011)). Simulated population data are then characterised by spatial and/or temporal metrics and conclusions are made about the landscape properties that better answer the ecological question at stake (*e.g.* Fahrig 1998).

For a given species, traits values —*e.g.* movement, reproduction and mortality rates—depend on environmental variables—such as amount and quality of feeding resources, (micro-)climatic conditions—that are spatially heterogeneous at the landscape scale (Honek et Jarosik 2000; Holland et Luff 2000). Traits sensitivity to environmental variables however varies from one species to another,

and from one trait to another (Jackson et Fahrig 2015). Spatial heterogeneity of the environmental variables will thus generate different population patterns according to their impacts on the species life-traits (Fahrig, Girard et al. 2015). When attempting to determine the effect of landscape structure through spatially explicit modelling approach, one has to include these processes both qualitatively (*i.e.* which traits have to be included and are they sensitive to environmental heterogeneity?) and quantitatively (what are the traits values with regard to the environmental variables?) (Schröder et Seppelt 2006). Unfortunately this knowledge is rarely avai-

lable in the literature, and hardly acquirable through experimentation, both because quantifying movement or reproduction is technically difficult, and that this quantification should be done along the whole environmental gradient. This calls for alternative methods to determine the traits-environment relationships.

While the traits-environmental variables relationships are responsible for the population patterns in space and time, it is reciprocally possible, under some conditions, to infer these processes from the analysis of population patterns (Fortin et Dale 2005 ; Grimm *et al.* 2005). In turn, population patterns are used as a signature of the underlying ecological process (McIntire et Fajardo 2009). This assumption is a central theme of movement ecology (Avgar *et al.* 2013 ; Nathan 2008) and some studies successfully inferred the movement processes underlying population patterns. McIntire, Rompré *et al.* (2013) for instance were able to determine the movement behaviour of a butterfly species by comparing population data from the field and data simulated by individual-based models carrying different mechanistic hypotheses.

Following the concept of "strong inference" initially formulated by Platt (1964), two conditions need to be fulfilled to ensure the robustness of the ecological processes inference : *i*) multiple competing hypotheses about the processes must be formulated *a priori* (Johnson et Omland 2004), and *ii*) the experimental design effectively allows to discriminate the competing hypotheses (McIntire et Fajardo 2009 ; Platt 1964). In the discipline of landscape ecology, where the experiment consists in sampling individuals in space and time, this second condition means that the sampling must be designed such as the population signatures associated to different processes are contrasted enough to be distinguishable (Schweiger *et al.* 2016 ; Schröder et Seppelt 2006). Using such a design on the field thus ensures that the extracted signature leads to the correct selection of the underlying ecological process.

Conversely, insufficient sampling may lead to a pattern description that is too coarse to disentangle the competing hypotheses formulated about the ecological processes (McIntire et Fajardo 2009). The question of optimal sampling design to unravel ecological processes from the population signature has thus been of an increasing interest for the last four decades (Mateu et Müller 2012). In this context, Zurell *et al.* (2010) emphasised the exciting potential of the use of virtual population data to evaluate methods for data sampling, naming this approach as the "virtual ecologist" methodology. However, to our knowledge, no methodological framework using simulated population data has been proposed to set up sampling over space and time with the specific objective of inferring the ecological processes driving a species dynamics from a set of *a priori* hypotheses. Defining a heuristic method based on the idea that ecological processes can be inferred from the population signature (Tilman et Kareiva 1997 ; Tuda 2007), we address this methodological issue.

After presenting the general approach of our framework, we illustrate its application through a real landscape based scenario to define *a priori* how sampling should be conducted to infer the ecological processes driving the population dynamics of a ground-dwelling carabid species, *Pterostichus melanarius*, among a set of competing hypotheses.

Although *P. melanarius* population distributions within the agroecosystem have been widely studied (*e.g.* Holland J. M 2004 ; C F George Thomas 2002), considerable gaps in our knowledge remain and must be filled if we are to manipulate the farm habitat in its favour (Holland et Luff 2000). For example, knowledge related to its preferred habitat types remains unclear, as well as the processes governing its heterogeneous distribution within and between fields (Holland et Luff 2000). Predicting population patterns in relation to landscape structure through an approach of landscape modelling cannot be conducted without such ecological knowledge (Jongejans *et al.* 2008 ; Schröder et Seppelt 2006).

We tackle this lack of knowledge about this key natural enemy species, by formulating competing mechanistic hypotheses that are susceptible to drive its population within the agricultural landscape. The goal of our study was to identify *a priori* sampling designs that would best allow selecting the correct ecological process at stake from the sampled data. In turn, we did not infer the ecological processes driving *P. melanarius* populations at the landscape scale. Rather we proposed a sampling design to do so in a real agricultural study area, within the range of sampling efforts usually made for studies where carabid are sampled within agricultural landscapes. The outputs of our study are *i)* a proposition of sampling design in terms of number of sampling points, number of sampling sessions in time, and number of landscape replicates in which to sample, *ii)* the associated classification model to be used for ecological process inference on the basis of sampled data and *iii)* the assessment of the population parameters uncertainty on the inference robustness, through a sensitivity analysis framework. Finally we discuss the implications of such framework, and what improvements remain to be done before being fully operational for ecologists that aim to infer ecological processes from the population signatures at the landscape scale.

2.2 Material and Methods

2.2.1 General heuristic approach

The general approach of our heuristic is illustrated on Fig. 2.1. It is designed to address the question of the optimal sampling strategies in order to maximise the ability to determine the biological-traits/environment relationships of the species of

interest. The approach is threefold. The first step consists of formulating competing ecological hypotheses from the available knowledge about the species of interest, translating them into mechanistic, spatially explicit population dynamics models (*Process Models PMs*), and simulating them on virtual landscapes. In the second step, sampling designs differing in their spatial and temporal components $\mathcal{S}_s, \mathcal{S}_t, \mathcal{S}_r$ (respectively the numbers of sampling locations, sampling dates and landscape replicates) are applied to sample the simulated population data. A set of summary statistics (hereafter named *signature* Φ) are computed on sampled data for each sampling design, defining the sampling dependent *PM* signatures. For each sampling design, the couples (PM, Φ) are used as inputs for a classification model that learns and predicts the *PM* as a function of their population signatures. The classification accuracy, defined as the rate of correct predictions of the *PMs* from their signatures, quantifies the sampling design quality : a high accuracy means that the sampling design generates *PM*-specific signatures that are easily distinguishable one from another. Using a classification model also allows to extract other information about the classification performances through the analysis of the confusion matrices for each sampling design. For instance, type I and II error rates (respectively "false positives" and "false negatives") for each *PM*, as well as their pairwise confusion, that is a proxy towards *PM* signatures similarity. Such information quantifies the uncertainty associated with the classification of ecological processes from sampled data.

Another valuable information is to describe the importance of the sampling components $(\mathcal{S}_s, \mathcal{S}_t, \mathcal{S}_r)$ on classification performances, and eventually predict the best way to distribute sampling components for any sampling effort $E = \mathcal{S}_s \cdot \mathcal{S}_t \cdot \mathcal{S}_r$. To do so, a smoothing model is used to fit the relationship between classification accuracy and sampling components.

The third step consists of assessing the effect of uncertainty of the *PMs* parameters on classification accuracy using a sensitivity analysis. This ranks the parameters impact and sheds light on the parameters that have to be more carefully calibrated to ensure the reliability of the sampling design selection.

Two points have to be accounted for when applying this framework to a real field study. First, the extent of the sampled landscapes has to be defined according to the available knowledge about dispersal abilities of the species of interest. Second, the maximum number of landscape replicates (\mathcal{S}_r), along with their properties in terms of configuration and composition and the variability of these properties have to be determined from the analysis of the real landscape. These landscape properties (extent, composition and configuration metrics) extracted from the landscape analysis determine the properties of the virtual landscapes upon which the *Process*

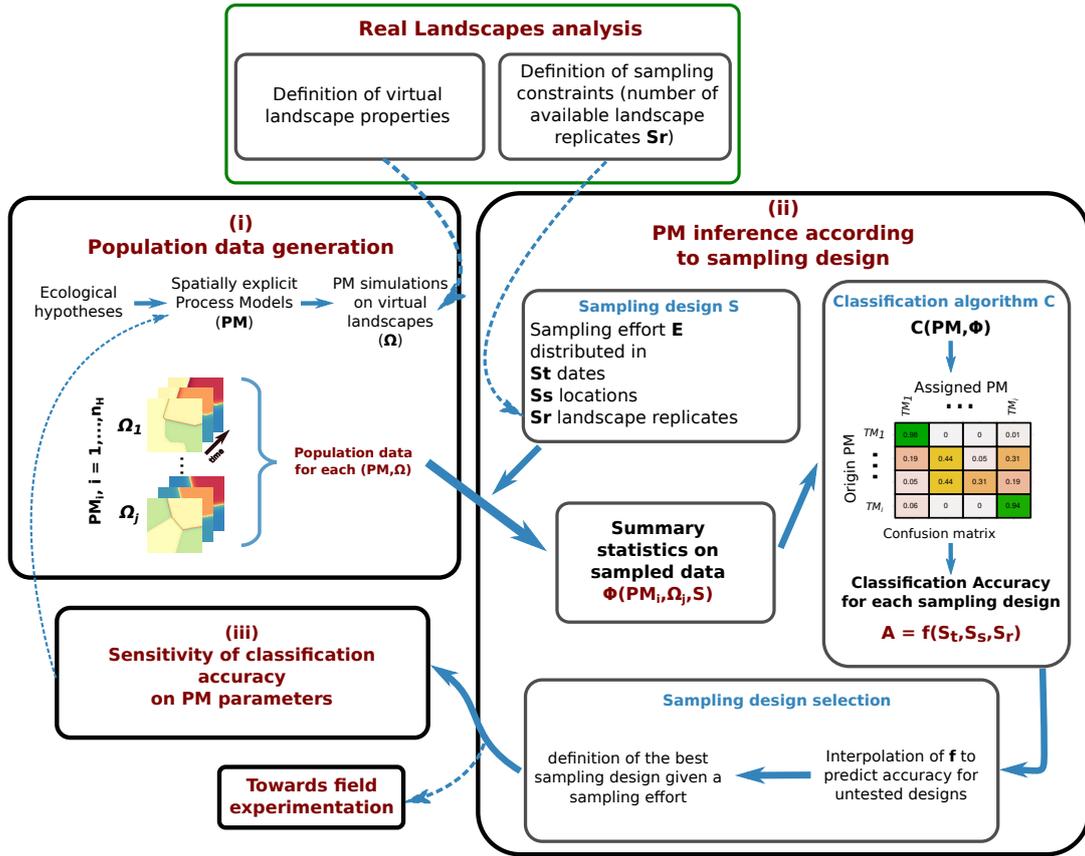


FIGURE 2.1 – Scheme of our heuristic framework. Step (i) generates virtual population data in space and time. In step (ii), these data are summarised for different sampling design and subject to a classification algorithm that infers the underlying process from these signatures. Classification accuracy is extracted for each classification procedure, interpolated as a function of sampling components. Once a sampling design is chosen, the sensitivity of PM parameters on classification accuracy is assessed in step (iii) to target the parameters that need proper calibration.

Models are simulated.

2.2.2 Definition of landscape model constraints and generation of virtual landscapes

Before designing the sampling experiment of any landscape ecology study, a preliminary analysis of the study area is necessary to select the landscape that are to be sampled (Gillespie *et al.* 2017). This selection process relies on the adequacy between the constraints induced by the ecological question(s) at stake and the constraints induced by the study area itself. Research-induced constraints might be constraints on the landscape structure, such as sampling landscapes presenting a gradient of mean patch size (*e.g.* Fahrig, Girard *et al.* 2015). Study area-induced constraints are the availability or not of landscapes whose characteristics match the research constraints.

Our study relies on the relationship between the ecological processes and the population dynamics signature they induce. We thus sought to limit the factors impacting population signatures that are not linked to differences in the underlying ecological processes. Both components of landscape heterogeneity, *i.e.* *compositional heterogeneity* (the amount and proportion of different habitats) and *configurational heterogeneity* (the way these habitats are arranged in space) are known to drive population patterns (Fahrig, Baudry *et al.* 2011). We sought to select landscapes presenting a low variability in these two factors.

In our example, we started from a real landscape of approximately 130 km^2 located in Brittany (north-western France) characterised by a SW-NE gradient of landscape structure—from hedged farmland, woodlands and permanent grasslands (high configurational heterogeneity), to open fields (low configurational heterogeneity; see Baudry *et al.* (2000) for more details). In 2011 (the season under study here), the area comprised around 75% agricultural surfaces, almost equally distributed between one third of grasslands, maize and winter wheat.

We extracted the structural properties in terms of compositional and configurational heterogeneity and included them as constraints into a landscape model to generate virtual landscapes sharing similar properties.

We set constraints for the landscapes selection such as *i)* the proportions of the different habitat types within a landscape were equal (*i.e.* 1/3 of the landscape surface) and *ii)* the value of configurational heterogeneity (*Edge density* (Edge Density, *ED*, a landscape-scale index of the interface length between different habitat types, expressed in m/ha ; (Cushman *et al.* 2008)) was the same within the landscapes. The landscape extent is defined by the area spanned by *P. melanarius* during its activity period (approximately $9.10^4 ha$; Alema 2014; Holland, Thomas *et al.* 2005).

Using floating windows of $300 \times 300 m$ with a translation step of 75 m in latitude and longitude, we scanned the study area, hereafter mentioned as ZAA (for Zone Atelier Armorique), to extract the landscape properties of habitat proportion and ED values. We retained landscapes containing more than 95% of agricultural covers, with the surface of maize, winter wheat and prairie equally distributed (*i.e.* $1/3 \pm 10\%$ for each cover type). Edge Density values vary from 38 m/ha to 228 m/ha within these filtered landscapes and the most common values of ED being about 192 m/ha ($\pm 10\%$).

We constructed a landscape model based on the tessellation of spatial point patterns to generate landscapes sharing the same properties in terms of composition and Edge Density values (Gaucherel 2008; Le Ber *et al.* 2009). We chose the family of Multitype-strauss HardCore (MSHC) point patterns, as they allow for multiple types of marks (land-cover types), hard-core properties (minimal distance between plot centroids), and both symmetrical positive and negative spatial dependence wi-

thin and between marks (Edelman 2012).

We generated 12 virtual landscapes, denoted $\Omega_{1,\dots,12}$, obtained by solving the following optimisation problem :

$$\min_{\nu} \left(\sqrt{(\widehat{ED}(\mathbf{X}(\nu)) - \overline{ED})^2} + \sum_{u=1}^U \sqrt{(\widehat{p}_u(\mathbf{X}(\nu)) - \overline{p}_u)^2} \right) \quad (2.1)$$

with U Number of land-cover types, \mathbf{X} Point pattern obtained from the MSHC Model, ν landscape model parameter vector, \widehat{p}_u Surface proportion of the land-cover type u in the generated landscape, \overline{ED} targeted Edge Density value, here 19.2 km.km⁻², \widehat{ED} Edge Density (km.km⁻²) value in the generated landscape, \overline{p}_u targeted surface proportion of the landcover type u , here 1/3 as equiproportion is constrained.

2.2.3 Process Models formulation and generation of population data

The reaction-diffusion formalism is suitable for modelling population dynamics with the possibility of accounting for spatial heterogeneity in the demography and/or diffusion parameters (Shigesada et Kawasaki 1997; Roques, Auger-Rozenberg et al. 2008). We used this formalism to formulate seven Process Models (*PMs*) with different responses to spatial heterogeneity.

Habitat quality stands for "*the resources and conditions present in an area that produce occupancy—including survival and reproduction—by a given organism*" (Hall et al. 1997). Habitat quality can be represented by the carrying capacity parameter, that implicetely accounts for factors such as amount and availability of resources, or (micro-)climatic conditions (Banks 1994; Barros et al. 2016). By giving it a spatially homogeneous—respectively heterogenous—value, one can represent habitat generalism—resp. habitat specialism—(e.g. Padrón et Trevisan 2006). We thus modelled the reaction component using the Verhulst logistic equation (see e.g. Banks 1994; Huzimura et Matsuyama 1999), possibly accounting for spatially heterogeneous habitat quality by setting a land-cover dependent value for the carrying capacity. The diffusion component was modelled by a Fokker-Planck equation for models with spatially heterogeneous diffusion, as it has proven robust for handling the population redistribution at habitat interfaces (Codling et al. 2008; Roques et Chekroun 2010; Bourhis, Poggi, Mammeri, Cortesero et al. 2015). For the models with spatially homogeneous diffusion, the Fokker-Planck diffusion simply writes as a Fick diffusion (*i.e.* the heat equation). The general formulation of the reaction-diffusion model reads :

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K(\mathbf{x}, t)} \right) + \Delta_{\mathbf{x}} (\mu(\mathbf{x}, t)P(\mathbf{x}, t)) \quad (2.2)$$

with \mathbf{x} standing for the two dimensional representation of space (x, y) , $P(\mathbf{x}, t)$ the population density at location \mathbf{x} and at time t , r the *per capita* growth rate of population, $K(\mathbf{x}, t)$ the carrying capacity in \mathbf{x} at time t , $\mu(\mathbf{x}, t)$ the diffusion parameter over \mathbf{x} and t , often called motility when applied to the Fokker-Planck diffusion (Alema 2014). Finally, $\Delta_{\mathbf{x}} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplace operator standing for the second derivatives in space, *i.e.* the diffusion. The *PM* formulations differ in the way their reaction and diffusion (hereafter *R* and *D*, respectively) parameters K and μ depend on the land-cover type. We defined four modalities for each of the *R* and *D* components :

- Homogeneous and Constant (*HC*) : the considered parameter (K or μ) does not depend on the land-cover type. Its value is set within the whole spatial domain and remains constant along with time.
- Inhomogeneous and Constant (*IC*) : the considered parameter (K and/or μ) depends on the land-cover type. Its value is a function of the land-cover type and it remains constant along with time.
- Homogeneous and time-varying (*HV*) : the considered parameter (K or μ) does not depend on the land-cover type. Its value is a decreasing exponential function of time.
- Null. (*N*) : there is no change in population size at the landscape scale or no population movements along the simulations.

By crossing the four reaction and four diffusion modalities and removing the non-spatial formulations, we formulated seven *PM* termed $\mathcal{M}_{iR,jD}$, $i, j \in \{N, HC, HV, IC\}$, each of which represents a hypothesis on the traits-environment relationship of the fictive carabid species. For the sake of parsimony, we chose not to represent an Inhomogeneous and time-varying modality, as it would imply defining a functional to describe the changes of the carrying capacity and diffusion parameters in time and within each land-cover, for which information for the calibration is too scarce.

Table 2.2 details the *PM* formulations. The parameter values are given in Table 2.3. The seven *PM*s were simulated on the 12 virtual landscapes $\Omega_{1,\dots,12}$ previously generated. Details on the numerical scheme used for the *PM* simulations are reported in Appendix B.2.

2.2.4 Effect of the sampling design on ecological process inference

Definition of the sampling designs

We defined a sampling design \mathcal{S} as the combination of three parameters : \mathcal{S}_t , the number of sampling sessions in time, \mathcal{S}_s , the number of sampling locations in space per landcover type and \mathcal{S}_r , the number of landscape replicates within which the population samplings are applied. The total sampling effort E is defined as the product of its spatial and temporal components $E = \mathcal{S}_t \cdot \mathcal{S}_s \cdot \mathcal{S}_r$. Aside from some exceptional sampling designs that involve a total effort of around 4000 points within a sampling campaign (*e.g.* Holland, Thomas *et al.* 2005), sampling designs involving carabid sampling rarely exceed around 1000 points (Bohan *et al.* 2011 ; Marrec *et al.* 2015). We defined values for \mathcal{S}_t , \mathcal{S}_s and \mathcal{S}_r so as the range of these "feasible efforts" is finely explored. We extended this range up to $E = 21600$ points to explore a potential asymptotic effect of sampling on *PM* classification. We thus defined the values of sampling parameters \mathcal{S}_t , \mathcal{S}_s and \mathcal{S}_r as follows :

$$\mathcal{S}_t = \{1, 2, 3, 5, 10\}$$

$$\mathcal{S}_s = \{1, 2, 3, 5, 10, 20, 30, 60\}$$

$$\mathcal{S}_r = \{5, 6, 7, 8, 9, 10, 11, 12\}$$

giving a total of 320 sampling designs to be tested. We sampled \mathcal{S}_s points at \mathcal{S}_t dates within the simulated population for each couple model-landscape (\mathcal{M}, Ω) . We assumed a random stratified sampling for the location of the \mathcal{S}_s sampling points within a landscape, *i.e.* \mathcal{S}_s points are randomly sampled within each of the three land-cover types. Sampling dates \mathcal{S}_t are regularly distributed within the simulation time. The \mathcal{S}_r landscape replicates were randomly taken out of the 12 Ω domains. We controlled random sampling so that every population returned by the *PMs* were sampled at the same locations and within the same landscapes. In order to take into account the stochasticity of the sampling locations, each sampling was replicated 10 times.

Performance assessment and sampling design selection

The idea of our work is to search for sampling designs that maximise the contrasts between the signatures so that the *PM* underlying a new data sample can be clas-

sified with a minimal risk of classification error.

Let y be the observation of a population data sample for which an underlying process PM is to be inferred from its descriptors Φ (a metrics vector describing the observation y). The issue of inferring the ecological process PM from a set of *a priori* hypotheses PM_m , given its signature Φ can be formalised as a classification issue. The problem consists in inputting Φ into a classification algorithm \mathcal{C} that outputs $P(y = PM_m|\Phi)$, which is the probability that the observation y , described by Φ , follows the ecological process PM_m . The selected process model \widehat{PM} for the dynamics y is then taken as the one with highest probability, *i.e.*

$$\widehat{y} = \arg \max P(y_i = PM_m|\Phi), \quad m = 1, \dots, M \quad (2.3)$$

The ability to correctly classify y thus depends on the specificity of the descriptors Φ with regard to the Process model. Formally, the classifier performances depend on the d -dimensional distances between the probability density functions (*pdf*) $f_m(\Phi)$, that are the estimated probability density of the descriptor Φ if the PM is m . When the signatures of two PM s are close, their *pdf* overlap and thus the PM s are hardly distinguishable.

For each sampling design \mathcal{S}_j , $j = 1, \dots, 320$, we fitted a classifier \mathcal{C}_j^* on the population observations (y^j, Φ^j) sampled with \mathcal{S}_j . The classification accuracy of the fitted classifier $A(\mathcal{C}_j^*)$ was assessed by a *K-fold* cross-validation procedure (Hastie, Tibshirani et Friedman 2009) :

$$A(\mathcal{C}_j^*) = \frac{1}{N_j} \sum_{k=1}^K \frac{\sum_{i=1}^{N_k} (\widehat{y}_{ik}^j = y_{ik}^j)}{N_k} \quad (2.4)$$

with $N_j = M \cdot \mathcal{S}_r(\mathcal{S}_j)$ the number of observations for the sampling design j , K the number of cross-validation folds, N_k the number of validation data observations in the fold k , \widehat{y}_{ik}^j the predicted PM for the validation observation, and y_{ik}^j the actual class of the validation observation.

The classifications were done using a multinomial logistic regression. This regression is solved with a gradient boosting algorithm which combines decision trees and boosting (*i.e.* combining many simple models to improve prediction). For details about the procedure, we refer the reader to Friedman (2001). We derived the classifiers using the R package `gbm` version 2.1-06.

The observation descriptor Φ is a vector of 15 descriptive statistics computed on the sampled population datasets. These descriptors are based on population means, variances and extrema (minimum or maximum population) aggregated along time and/or space and are thus global descriptors of the population at the landscape scale all over the sampling season. While finer population descriptors could have been more relevant to make it easier the classification of the ecological processes, we chose these metrics as they are conceptually simple and mimic the data pooling in space and time usually encountered in the literature involving carabid sampling. Tables B.1 and B.2 in Appendix B.3 detail the formulation of population metrics. To further investigate the PM signatures similarity in relation with sampling design, we also computed the pairwise confusion rate, which is the sum of type I error (incorrectly classifying PM_i as $PM_m, i \neq m$) and type II error (incorrectly classifying PM_m as $PM_i, i \neq m$).

We selected the best sampling design \mathcal{S}^* as the one that maximised the classification accuracy within the range of sampling efforts usually encountered in the literature ($E \leq 1000$, 200 sampling designs). The classification model associated with this sampling design is thereafter referred to as the *reference classifier* \mathcal{C}^* .

Relative importance of sampling parameters on classification accuracy

Owing to the large number of possible combinations for the sampling parameter values, it might not be possible to test every sampling design for classification accuracy assessment. Smoothing models can be used to interpolate classification accuracies of untested designs, and quantify the relative importance of the sampling parameters.

We used General Additive Models (GAMs) (Hastie et Tibshirani 1990) to fit the relationship between the sampling design parameters (and possibly their interactions) and classification accuracy

$$g(\mathcal{A}) = f_1(\mathcal{S}_t) + f_2(\mathcal{S}_s) + f_3(\mathcal{S}_r) + f_4(\mathcal{S}_t, \mathcal{S}_s) + \dots + f_i(\mathcal{S}_t, \mathcal{S}_s, \mathcal{S}_r) \quad (2.5)$$

where f_i are smoothing functions of the variables \mathcal{S}_k , $k = s, t$, or r . We first fitted all the possible models to be formulated from the sampling variables \mathcal{S}_t , \mathcal{S}_s , \mathcal{S}_r and their 2^{nd} and 3^{rd} order interactions with data. Smoothing function estimation was performed using penalised thin plate regression splines (Wood 2003). We then selected the model g^* as the one presenting the lowest Akaike's Information Criterion (AIC) (Johnson et Omland 2004).

2.2.5 Effect of model calibration uncertainty on species traits inference

The reaction and diffusion parameter values of the *PMs* were derived from the literature and/or expert knowledge and thus are likely to diverge from the values that govern population dynamics on the field. Consequently, if a certain ecological process drives the species population in the field, but with values differing from the ones we set for the a priori hypotheses, the reference classifier might not be able to recognise its signature from the sampled data, and mistake it for another process. We wanted to assess the effect of an error in the parameter values on the classification accuracies.

Let v_1^*, \dots, v_p^* be the reference parameter values we set by expert knowledge and/or literature, and v_1', \dots, v_p' the parameter vector that would optimally describe the population dynamics on the field. Let $y^* = PM_m(v_1^*, \dots, v_p^*)$ and $y' = PM_m(v_1', \dots, v_p')$ be the vectors combining the sampled observations of the population dynamics generated by these models, and Φ^* and Φ' their respective descriptors. We want to assess the effect of the divergence between v^* and v' on the loss of classification performances, *i.e.* the risk of misclassification due to a wrong initial calibration of the Process Models. Formally, we conducted a global sensitivity analysis of the PM parameters v_1, \dots, v_p on the loss of classification accuracy.

We used the Morris method (Morris 1991) to design the parameter space matrix V' , using a one-at-a-time factorial design with 10 levels for each parameter, a grid jump δ of $10/2 = 5$, and $H = 30$ trajectories, leading to a table of $n = 420$ parameter sets. We simulated the process models and extracted the associated population data, and calculated the loss of classification accuracy for each parameter set. We computed the absolute mean (μ^*) and variance (σ) of the parameters elementary effects on the loss of classification accuracy to assess the parameters sensitivity. Details of the procedure is given in Appendix B.4.

2.3 Results

2.3.1 Effect of sampling design

The ability to distinguish the ecological hypotheses from each other, approached by the classification accuracy, strongly relies on the way sampling is conducted. Accuracy can be approximated as a logarithmic function of total sampling effort ($R^2 = 0.84$), thus showing a dramatic increase within the range of realistic sampling efforts to reach a quasi asymptotic behaviour (fig 2.4). At the same time, variability of accuracy per sampling design decreases with the logarithm of total sampling

effort ($R^2 = 0.65$). On the global trend, increasing sampling effort thus improves mean classification accuracy and reduces its variability. For a given sampling effort, different sampling designs lead to different mean accuracies, especially for weak sampling efforts. As an example, for an effort of $E = 300$, possibly reached with seven sampling designs, mean accuracy ranges from 0.69 to 0.85. Such variability shows the importance of wisely distributing sampling effort E between its components \mathcal{S}_t , \mathcal{S}_s and \mathcal{S}_r . On the other hand, different ways of managing the effort might provide equivalent classification performances. Given the same example of sampling effort, two sampling strategies allow getting an accuracy of $\mathcal{A} = 0.85$, either taking $(\mathcal{S}_t = 2, \mathcal{S}_s = 10, \mathcal{S}_r = 5)$ or $(\mathcal{S}_t = 2, \mathcal{S}_s = 5, \mathcal{S}_r = 10)$. Getting a similar accuracy with different designs allow choosing the sampling strategy that best fulfils the sampling campaign constraints (*e.g.* number of available landscape replicates, or possible sampling dates).

The GAM model used to fit the relationship between the sampling parameters and the classification accuracy explained 82.8% of the deviance. The three sampling parameters, including their 2nd and 3rd order interaction have a significant impact on classification accuracy. All the predicted accuracies (red dots on fig. 2.4) all fall within the range of "observed" accuracies (blue error bars on fig. 2.4), ensuring its reliability to interpolate accuracies for untested designs.

The relative influence of the sampling parameters on the accuracy depends on the location in the 3-dimensional space $(\mathcal{S}_t, \mathcal{S}_s, \mathcal{S}_r)$, as observable on fig. 2.5. In turn, the sampling parameter to be adjusted for a better classification depends on the current location of the sampling design in the parameter space. Interestingly, a change in parameter value does not necessarily means an increase in its value. For instance, for a given value of \mathcal{S}_s and/or \mathcal{S}_r , an increase of \mathcal{S}_t from two to three results in a decrease in classification accuracy. Provided the impacts of sampling parameters on accuracy are non-linear and interacting, there is no rule-of-thumb for systematically prioritising a dimension rather than another. However, optimal sampling designs with regards to user-defined sampling constraints can then be obtained by searching for local maxima of the fitted GAM model.

Some ecological processes generate close population signatures with each other, making them harder to distinguish with each other. PM pairwise-confusion rate is a finer descriptor of classification performances. It is calculated as the sum of type I error (incorrectly classifying model A as another model) and type II error (incorrectly classifying another model as model A). It thus provides an information about how PM signatures are close to each other : two PM exhibiting a high confusion have close signatures, as the classification algorithm is not able to discriminate one from another. Fig. 2.6 show the pairwise distributions of confusion rate within the range of fea-

sible efforts ($E \leq 1000$). The higher the confusion, the most often are the two models mistaken one for another for a given sampling design. For most of the PM couples, confusion is null or weak within the feasible sampling effort, meaning that they are never mistaken one for another, or that when they are, the error is low. Three couples however show non-negligible confusion rates : $(PM_{NULL-IC}, PM_{HC-IC})$, $(PM_{NULL-IC}, PM_{IC-IC})$ and (PM_{IC-HC}, PM_{IC-HV}) . This means their population signature are so close that their underlying processes are hardly distinguishable, *i.e.* that the sampling design must be selected carefully to disentangle the associated hypotheses. Confusion rates between the PM (PM_{IC-HC}, PM_{IC-HV}) never equal zero , meaning that regardless of the sampling design within the range of feasible efforts, these two processes can never be properly disentangled.

FIGURE 2.2 – Ecological hypothesis for the reaction and diffusion components, model names, model formulations and involved parameters. HC : Homogeneous Constant, HV : Homogeneous Varying, IC : Inhomogeneous Constant.

Reaction component	Diffusion component	Model name	Formulation $\frac{\partial P(\mathbf{x}, t)}{\partial t} =$	Parameters involved
NULL	IC	$\mathcal{M}_{NR,ICD}$	$\Delta_x (\mu(\mathbf{x}, t)P(\mathbf{x}, t))$	μ_1, μ_2, μ_3
HC	IC	$\mathcal{M}_{HCR,ICD}$	$rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K}\right) + \Delta_x (\mu(\mathbf{x}, t)P(\mathbf{x}, t))$	$r, K, \mu_1, \mu_2, \mu_3$
HV	IC	$\mathcal{M}_{HVR,ICD}$	$rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K_0 e^{-\beta_r t}}\right) + \Delta_x (\mu(\mathbf{x}, t)P(\mathbf{x}, t))$	$r, K_0, \beta_r, \mu_1, \mu_2, \mu_3$
IC	NULL	$\mathcal{M}_{ICR,ND}$	$rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K(\mathbf{x})}\right)$	r, K_1, K_2, K_3
IC	HC	$\mathcal{M}_{ICR,HCD}$	$rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K(\mathbf{x})}\right) + \mu \Delta_x P(\mathbf{x}, t)$	r, K_1, K_2, K_3, μ
IC	HV	$\mathcal{M}_{ICR,HVD}$	$rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K(\mathbf{x})}\right) + (\mu_0 e^{-\beta_d t}) \Delta_x P(\mathbf{x}, t)$	$r, K_1, K_2, K_3, \mu_0, \beta_d$
IC	IC	$\mathcal{M}_{ICR,ICD}$	$rP(\mathbf{x}, t) \left(1 - \frac{P(\mathbf{x}, t)}{K(\mathbf{x})}\right) + \Delta_x (\mu(\mathbf{x}, t)P(\mathbf{x}, t))$	$r, K_1, K_2, K_3, \mu_1, \mu_2, \mu_3$

FIGURE 2.3 – Process models parameter values for the sampling design selection (reference value) and ranges for the Global sensitivity analysis (§2.3.2)

Name	Description	Number of models	Names	Reference value	GSA parameter range	Method reference
r	<i>per capita</i> growth rate	6	all except $\mathcal{M}_{NR,ICD}$	4.82	2.41-7.22 (reference value \pm 50%)	Retho et al. (2008) Alema (2014)
K_1, K_2, K_3	Normalised Carrying capacity in land-cover types of "mid", "high" and "low" quality, respectively	4	$\mathcal{M}_{ICr,ICd}$; $\mathcal{M}_{ICr,HCd}$; $\mathcal{M}_{ICr,HVd}$; $\mathcal{M}_{ICr,ICd}$	$K_1 : 0.55$ $K_2 : 1$ $K_3 : 0.1$	$K_1 : 0.35 - 0.65$ $K_2 : 0.8 - 1$ $K_3 : 0.05 - 0.2$	Nisbet and Gurney (1976) Banks (1994)
K	Carrying capacity in all land-cover types	1	$\mathcal{M}_{HCr,ICD}$	0.55	0.3 - 0.8	cf. Appendix .1
K_0	Initial carrying capacity	1	$\mathcal{M}_{HCr,ICd}$	0.9	0.8- 1	cf. Appendix .1
k_r	Exponential decrease parameter for carrying capacity	1	$\mathcal{M}_{HCr,ICd}$	1.11	1.55 - 1.66	cf. Appendix .1
μ_1, μ_2, μ_3	Diffusion parameter in land-cover types of "mid", "high" and "low" quality, respectively	4	$\mathcal{M}_{NR,ICd}$; $\mathcal{M}_{HCr,ICd}$; $\mathcal{M}_{HCr,ICd}$; $\mathcal{M}_{ICr,ICd}$	$\mu_1 : 1.51 \cdot 10^{-2} t^{-1}$ $\mu_2 : 5.29 \cdot 10^{-3} t^{-1}$ $\mu_3 : 2.5 \cdot 10^{-2} t^{-1}$	$\mu_1 : 1.06 \cdot 10^{-2} - 1.96 \cdot 10^{-2} t^{-1}$ $\mu_2 : 7.9 \cdot 10^{-4} - 9.8 \cdot 10^{-3} t^{-1}$ $\mu_3 : 2.05 \cdot 10^{-2} - 3.0 \cdot 10^{-2} t^{-1}$	Alema (2014)
μ	Diffusion parameter in all land-cover types	1	$\mathcal{M}_{ICr,HCd}$	$1.51 \cdot 10^{-2} t^{-1}$	$1.5 \cdot 10^{-2} - 3.0 \cdot 10^{-2} t^{-1}$	cf. Appendix .1
μ_0	Initial diffusion parameter	1	$\mathcal{M}_{HCr,ICd}$	$3.46 \cdot 10^{-2}$	$1.73 \cdot 10^{-2} - 5.19 \cdot 10^{-2}$	cf. Appendix .1
μ_m	Exponential decrease parameter for diffusion parameter	1	$\mathcal{M}_{ICr,HVd}$	1.97	0.98 - 2.95	cf. Appendix .1

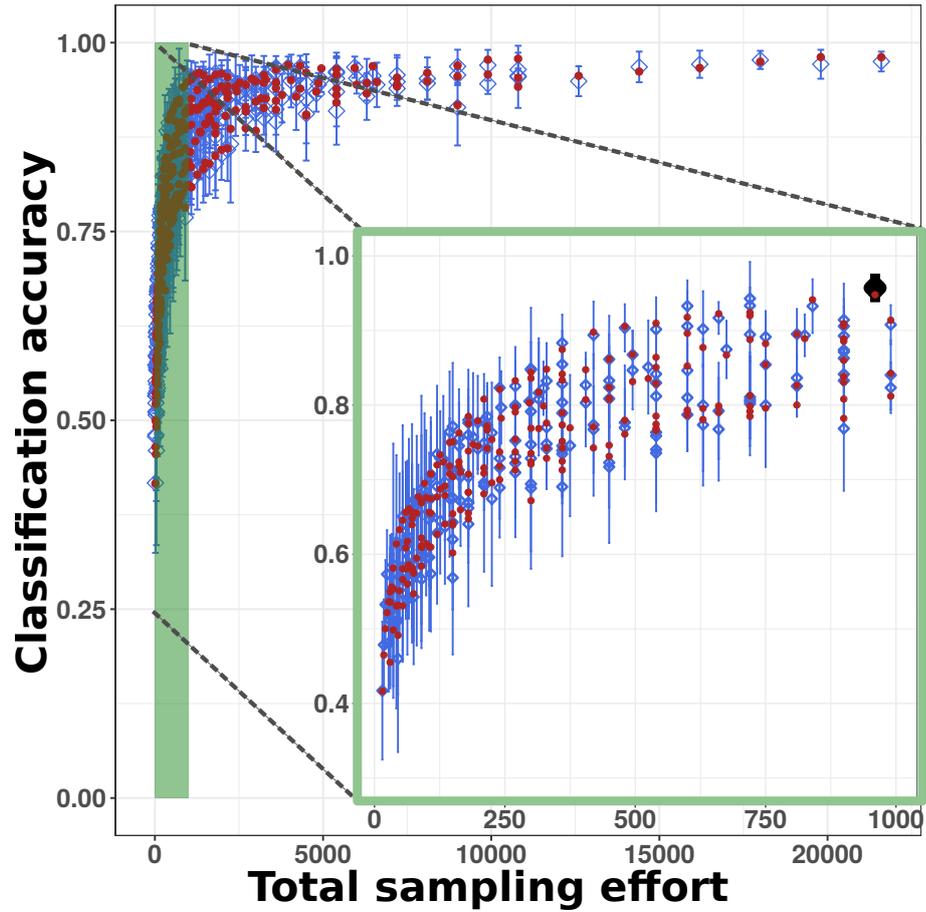


FIGURE 2.4 – Classification performances according to the total sampling effort ($E = S_t \cdot 3S_s \cdot S_r$). Blue diamonds with error bars represent the mean and standard deviation performance of "observed" data ($n=10$ replicates per sampling design). Black diamond with error bar represents the selected design S^* . Red dots are the predicted performances using the GAM model for the same sampling designs. The green area highlights the range of usually encountered sampling effort in studies at the landscape scale within the explored literature.

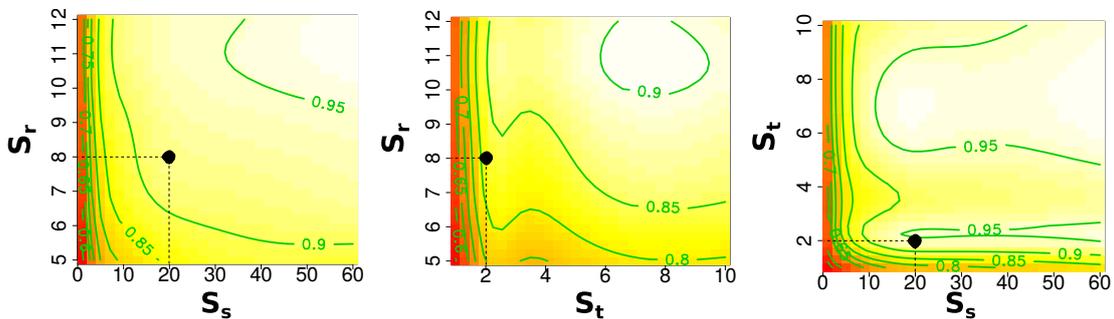


FIGURE 2.5 – Contour plots of the classification accuracies predicted by the fitted GAM as a function of sampling components S_t , S_s and S_r pairwise. The black dot corresponds to the location of the selected sampling design S^* . Notice that this selected design correspond to a three-dimensional optimum and, therefore, the black dots does not overlap with the optima of the two-dimensional contours figured here.

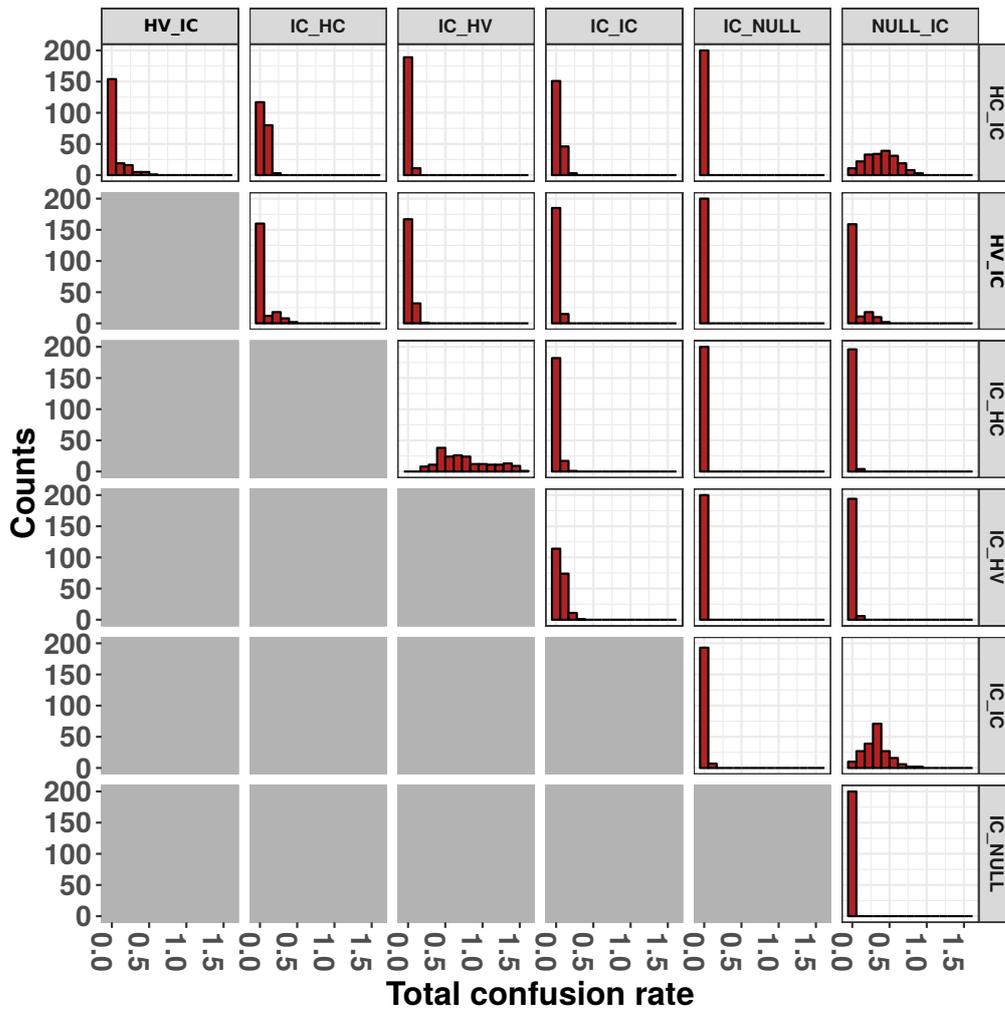


FIGURE 2.6 – Pairwise confusion between models within the range of feasible designs (200 sampling designs). Each panel represents the distribution of the confusion rates between the PMs in rows and the PMs in columns.

After this global analysis, we focused on the most performing sampling design \mathcal{S}^* (mean accuracy $\mathcal{A} = 0.95$, $sd=0.019$, orange diamond on fig. 2.4) within the feasible sampling effort range. The total sampling effort is 960 points, distributed in $\mathcal{S}_t = 2$, $\mathcal{S}_s = 20$, and $\mathcal{S}_r = 8$. Its associated confusion table (Fig. 2.7) is the tool to interpret the classification results once field data will be obtained, summarised and be used as input data for the associated classification model. The matrix diagonal gives the probability that the predicted model is the actual model governing population dynamics on the field (*i.e.* "true positives"). For instance, if the classifier interprets the field signature as being originated by the $IC - HC$ process (heterogeneous constant reaction, homogeneous diffusion), the probability that the real process is indeed $IC - HC$ is 0.84, and the probability to be mistaken for the $IC - HV$ process (same reaction, decreasing dispersal with time) is 0.16. Few or no confusion (type I and II errors $< 5\%$) are associated with the five other PM.

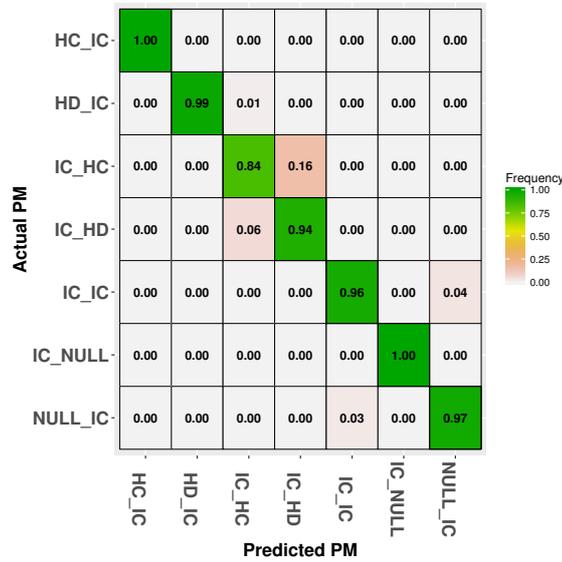


FIGURE 2.7 – Confusion matrix associated with the selected sampling design.

2.3.2 Sensitivity analysis

Correct calibration appears to be crucial with regards to classification accuracy. Loss of classification accuracies showed a wide range of values along the exploration of the PM parameters space given by the Morris plan, ranging from 0.02 to 1, with a mean 0.52 and standard deviation 0.20. All parameters have a significant impact on the loss of classification accuracy (Fig. 2.8). Parameters with high μ^* values are the most important regarding linear and additive effects on the loss of classification accuracy, whereas the ones with high σ values impact it non-linearly and/or with interactions with other parameters (Morris 1991). Parameters sensitivity on the classification accuracy can be interpreted as their impact on the population dynamics. A high sensitivity of a parameter means that a change in its value generates a population signature that is so different from the ones generated by the initial hypotheses that the reference classifier is not able to recognise it. Regarding this, K_3 the carrying capacity defining the value of "good quality" habitats, appears to be the most impacting parameter on population dynamics. As it determines the carrying capacity of the largest reservoir of population, it strongly drives the densities computed at the landscape scale. μ_2 , the diffusion value in low quality habitats (*i.e.* the high value of the diffusion ability), also appears to have a strong impact on classification accuracy and, as an underlying cause, on population dynamics. This highlights the importance of the organism ability to leave the poor quality habitats on the population signature at the landscape scale.

The impact of these two parameters highlights the importance of acquiring a good knowledge, either through literature or by conducting small scale abundance sampling, about *i)* what could be the good quality habitat in the landscape, and to correctly quantify its carrying capacity, and *ii)* what is the maximum diffusion ability of the organism within the landscape.

The type I error rate provides information about how processes are mistaken one for another. A type I error of *e.g.* 0.2 between PM_i and PM_j means there is a probability of 0.2 that the model PM_i inferred from the field signature is actually the model PM_j . Thus, using such classification performance metric is a useful tool to reduce the number of ecological hypotheses formulated a priori, accounting for the potential error in parameter values. Indeed, if a model PM_i gets a 0 type I error with model PM_j , regardless of the parameter values, then PM_j can be removed from the set of ecological hypotheses governing the species dynamics. Reading Fig. 2.9 row-wise, we can thus see that if PM_{HC-IC} is inferred from the field signature, there is no chance that the actual process is PM_{IC-HC} , PM_{IC-HV} or $PM_{IC-NULL}$, regardless of the error associated with the parameter calibration. The number of hypotheses is thus reduced from seven to four. A further analysis of the confusion

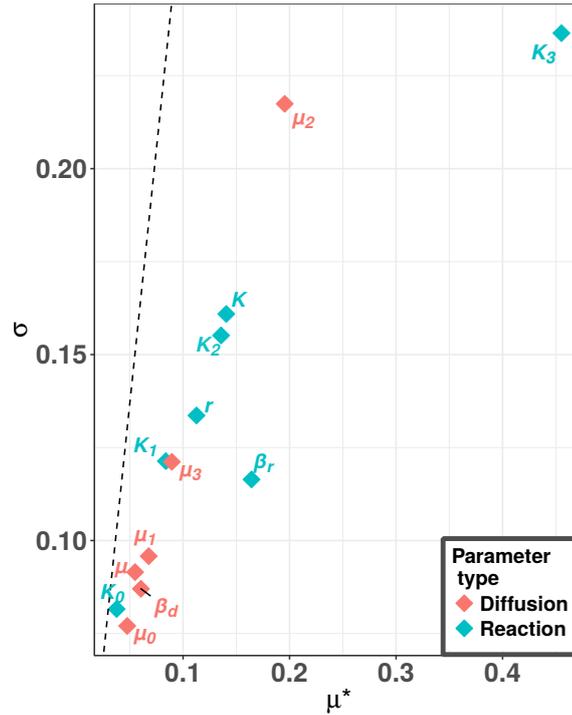


FIGURE 2.8 – Absolute mean and standard deviation of parameters Elementary Effects for the sensitivity analysis on reference classifier and simulated "field" data. Parameters lying below the dashed line ($\mu^* = 2SEM$ with $SEM = \sigma/\sqrt{H}$ and $H = 30$ being the number of trajectories, SEM for Standard Error of the Mean) have a statistically significant impact on the model output with a 95% confidence interval. Red and blue parameters are linked to the reaction and diffusion components, respectively.

matrices would allow quantifying the risk of error associated with classifying model $HC - IC$ as any of the remaining hypotheses. This would be a relevant extension of our framework.

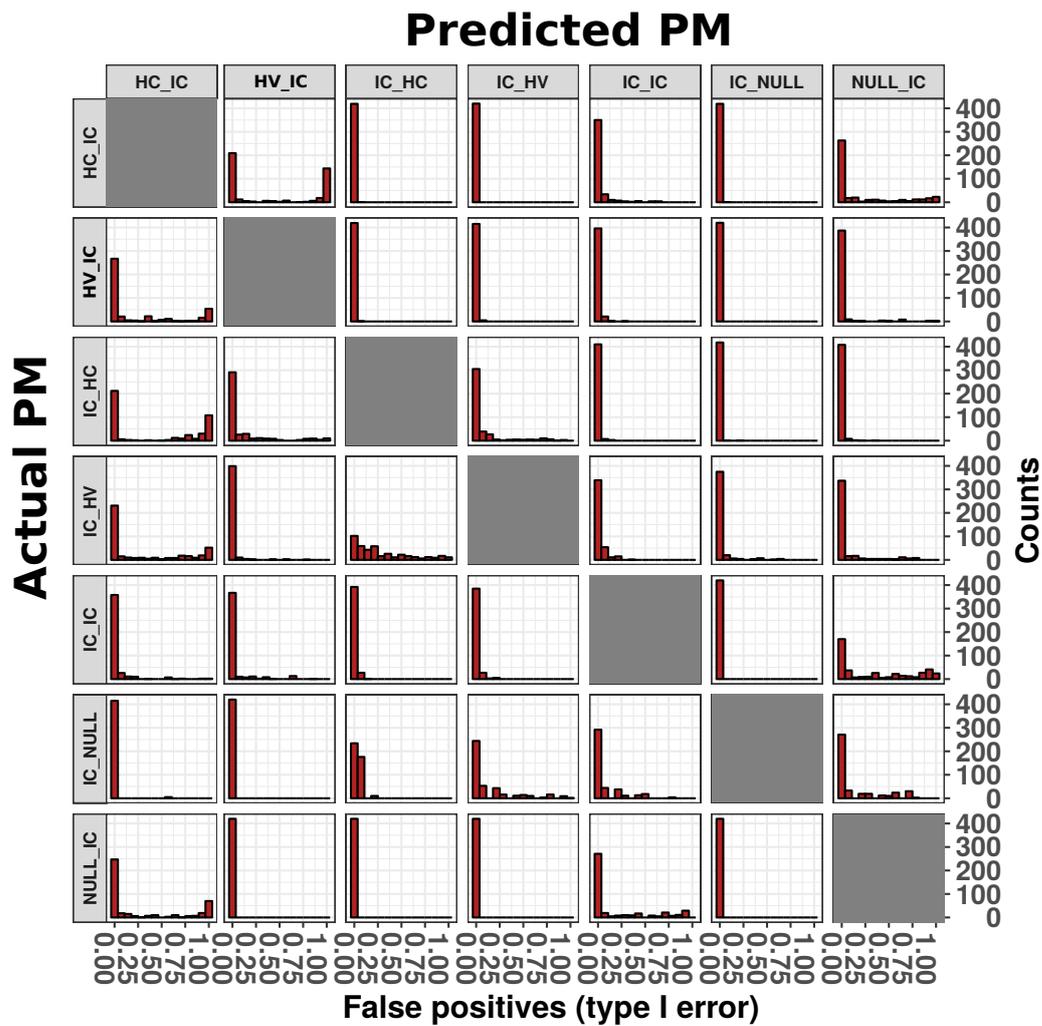


FIGURE 2.9 – Distributions of pairwise type I errors (incorrect classification of the *PM* in row as the *PM* in column).

2.4 Discussion and conclusions

2.4.1 Summary and implications

Incorporating mechanistic processes in population dynamics models is a crucial step for landscape-based species-management (Jongejans *et al.* 2008 ; Vinatier *et al.* 2011). Hence we developed a heuristic to select sampling strategies for processes inference on the basis of the analysis of population signature in space and time. Sampling quality is defined by its ability to extract the population signature so that it can be unambiguously related to a single ecological process. This was approached by assessing the performances of a classification model fitted on the sampled data. The higher the performances, the easier the process-specific signatures are distinguished by the classification model.

The variability of classification performances within the range of usually encountered sampling effort underlines the importance of determining *a priori* the way to distribute the effort between its spatial and temporal components. For a given effort, several designs can bring similar classification performances. The existence of alternative sampling strategies has valuable practical implications as it allows to choose the strategy that best matches the field constraints (*e.g.* number of possible visits on the field or number of available landscape replicates on the study area).

Using a smoothing model to fit the accuracy as a function of sampling components both allows to interpolate classification performances of untested designs, and to explore more finely the relative importance of sampling parameters. This showed that the relationships between the sampling parameters and classification accuracy are both non-linear and interacting. The relative importance of the sampling parameters thus depends on the location in the parameter space ($\mathcal{S}_t, \mathcal{S}_s, \mathcal{S}_r$) and no rule-of-thumb about sampling strategy can be defined. However, local maxima of classification accuracy are straightforwardly available when constraints about the sampling components are set through the use of the interpolation model.

While classification accuracy is a global proxy towards the ability to discriminate the ecological hypotheses, it does not give information about the signature similarity arising from different processes. More information can be extracted from the confusion matrices.

Pairwise confusion rates (sum of type I and II errors) quantify the similarity between the processes regarding their population signatures. Frequent errors between two models show that their signatures are so close that they can hardly be distinguished by the classification model. In our case, given the PM parameter values, the sampling effort usually made, the landscape properties and the population metrics used, two PMs ($IC - HC$ and $IC - HD$) could never be totally discriminated (*i.e.* non-null confusion rate regardless of the sampling design). It is interesting to show

that distinct ecological processes may lead to similar population patterns. This is interesting *per se* as it allows to make a typology of the processes on the basis of their pattern similarity (through *e.g.* hierarchical clustering). Moreover, assessing the steadiness of these patterns similarity along a gradient of landscape properties would then provide precious insights for landscape-based species management. Indeed, species driven by distinct ecological processes but that generate close patterns regardless of the landscape properties would be equally sensitive to landscape management strategies. Thus one or another process model could be indifferently used to explore new landscape scenarios, without necessity to get the finer information about the underlying ecological process.

Once we selected an example of optimal sampling design, we questioned the sensitivity of the classification accuracy on the *PM* parameters. In turn, we showed the effect of the variation in *PM* parameters values on the loss in classification accuracy compared to the one obtained with the reference parameter set. By doing so, we indirectly showed the importance of the parameters in the formation of the population patterns : for a sensitive parameter, a discrepancy between its reference value and its actual value in the field will generate a discrepancy in the patterns such as the classifier is not able any more to identify the underlying process. Sensitivity analysis is thus an important step, as it discerns the parameters that need proper calibration because of their strong influence on patterns population patterns. Uncertainty on the sensitive parameters needs to be quantified and if necessary reduced through *e.g.* lab or microcosm experimentation (Alema 2014).

We also showed that extracting finer classification metrics allows better qualification of the *PM* signature similarities, accounting for the uncertainty associated with the parameter values. Here again, the risk of mistaking a process for another can be quantified, and the number of competing ecological hypotheses can thus be reduced.

2.4.2 Working prospects and conclusion

We built a heuristic allowing to determine *how much* to sample and how to distribute it between its spatial and temporal components if one is to tackle the issue of inferring the ecological processes driving a species dynamics. As our method relies on process classification from the population signatures, any way to better capture contrasts between signatures originating from different processes is likely to increase classification performances. We identified three ways to do so.

First, *PM* signatures could be described at finer temporal and spatial scales. Instead of using data pooled at the landscape scale and over the whole season, population metrics could be computed per land-cover type and per sampling date. It is likely that these finer variables are more able to highlight the contrasts between

the population signatures. Second, one could take advantage of the "edge effect". Indeed, species experiencing spatial heterogeneity in their traits values (*e.g.* dispersal ability, mortality varying according to the landcover) exhibit specific signatures at the interface between two habitats (J. Haynes et T. Cronin 2006 ; Blitzer *et al.* 2012). Accounting for the distance from edges when designing sampling might ease the discrimination between homogeneous and heterogeneous processes. In a preliminary study (unpublished works), we have shown that for identical sampling designs, classification accuracy increases with *edge density* values. Whenever possible, we suggest to conduct sampling within landscapes presenting high *edge density* values if the objective is to infer the ecological mechanisms underlying the population distribution. Finally, the framework could greatly benefit from a method suggesting *when* and *where* to sample. By computing the distances between the populations densities $P(\mathbf{x}, t)$ originating from the different processes, it would be possible to target locations in (\mathbf{x}, t) that maximise the contrasts between the population signatures, thus increasing the classification performances.

The major constraint of our heuristic is that it require a certain amount of knowledge *a priori* about the species of interest to define i) the a priori competing ecological hypotheses and ii) the range of possible values for the Process Model parameters. Moreover, it is not possible to ensure that the initial hypotheses include the process that is at stake in the field. This is generally the case with hypotheses-driven inductive studies : we start from the initial set of ecological hypotheses and assess whether one of them correctly explains the field data. If not, *i.e.* if the fitted classifier does not manage to unambiguously classify the process on the basis of the field signature, the ecological hypotheses are refined or new ones emerge. This heuristic is thus an iterative process between field and model, where each step allows removing unsatisfactory hypotheses and build new ones that are more likely to explain observations (Platt 1964 ; Johnson et Omland 2004). In that sense, our framework is applicable to any species for which i) there exist competing hypotheses about the processes driving the observed population patterns (implying that population patterns for this species have been at least partially described) and ii) there exist a minimal knowledge (from experiment, literature or expert-based) about the species such as ranges for the parameter values of the ecological process(es) of interest can be set.

We are aware that this heuristic is not readily operational for field ecologists, as it still requires modelling skills and is time consuming. The shortest term perspective of this work is to render it readily usable by releasing an adaptive and open source package. We claim that such modelling framework helping to design sampling strategies prior to field experiments should be the rule rather than the exception. In addition to be cost-saving, such approach provides a design that is specifically

adapted to the ecological question at stake. Improving sampling leads to a better knowledge about the processes, which in turn allows a more thorough formulation of mechanistic population models, that is at last the necessary condition to build robust forecasting about the effect of landscape structure on population dynamics. Hence such framework should contribute to a better prediction of species distributions under the exploration of landscape scenarios, and thus provide a sounder basis for landscape-based species-management studies.

Acknowledgements

B. Bellot's Ph.D. was supported by a grant from the Ministry of Higher Education and Research. The authors are also grateful to Thomas Guyet and Fabrice Vinatier for their wise methodological advices and redaction inputs. We are also very grateful for the relevant insights provided by the two anonymous reviewers.

3

Perspectives d'amélioration concernant le mode d'échantillonnage

La méthode heuristique que nous avons présentée dans cet article vise à déterminer *a priori* des stratégies d'échantillonnage qui maximisent la robustesse de l'inférence des processus écologiques. Nous définissons une bonne stratégie d'échantillonnage comme permettant de maximiser les performances d'un algorithme de classification des modèles de processus sur la base de leurs motifs. Toute méthode permettant de capturer au mieux les contrastes entre les dynamiques de population est donc susceptible d'améliorer les performances de classification et donc la robustesse de l'inférence des processus.

Parmi les méthodes permettant de capturer au mieux les contrastes entre les motifs associés aux différents processus écologiques (cf. 2.4.2), nous avons évoqué l'amélioration des modalités d'échantillonnage. Dans les travaux présentés, nous avons fait le choix d'un échantillonnage aléatoire stratifié par type d'occupation du sol. Si ce choix est parcimonieux et représente une pratique courante en écologie du paysage, il est très probable que les performances de classification bénéficient d'un échantillonnage plus ciblé dans l'espace et le temps. En effet, comme la classification est basée sur les distances entre les descripteurs de population, échantillonner en des points de l'espace-temps qui maximisent ces distances est susceptible d'induire de meilleures performances de classification que pour un échantillonnage aléatoire.

Une voie d'amélioration pertinente des travaux que nous avons présentés serait donc de quantifier en tout point de l'espace et du temps la distance entre les densités de populations issues des différents processus écologiques et de concentrer l'effort d'échantillonnage sur les points-dates associés à un contraste maximum. Suivre cette démarche permettrait de plus d'utiliser directement les densités de population et non des métriques agrégées dans le temps et dans l'espace qui, comme nous l'avons discuté précédemment, ont tendance à réduire les contrastes entre les motifs de population.

En guise de travaux préliminaires pour cette voie d'amélioration, nous avons extrait une métrique pour quantifier ce contraste à partir des données de population virtuelles, appelée la distance de Gibbs, donnée par :

$$G = \sum_{i=1}^{M-1} \sum_{j=i+1}^M \frac{1}{(d_{ij})^2} \quad (3.1)$$

avec M le nombre de modèles de processus, d_{ij} la distance entre la densité de population en $\{x, y, t\}$ associée au modèle i et celle du modèle j , soit $d_{ij} = P_i(x, y, t) - P_j(x, y, t)$. Cette métrique étant basée sur l'inverse de la distance entre les densités de population, elle prend une faible valeur là où les contrastes sont élevés. La figure 3.1 montre les densités de populations issues de 4 modèles de processus à un temps donné (la fin de la simulation), et la distance de Gibbs issue de ces données en tout point de l'espace. Comme nous le supposons dans les perspectives de l'article (2.4.2), il semble exister un effet des interfaces sur les contrastes entre les modèles, et échantillonner à proximité des interfaces (faible valeur de G) semble être pertinent si l'on souhaite inférer les processus écologiques à partir des observations.

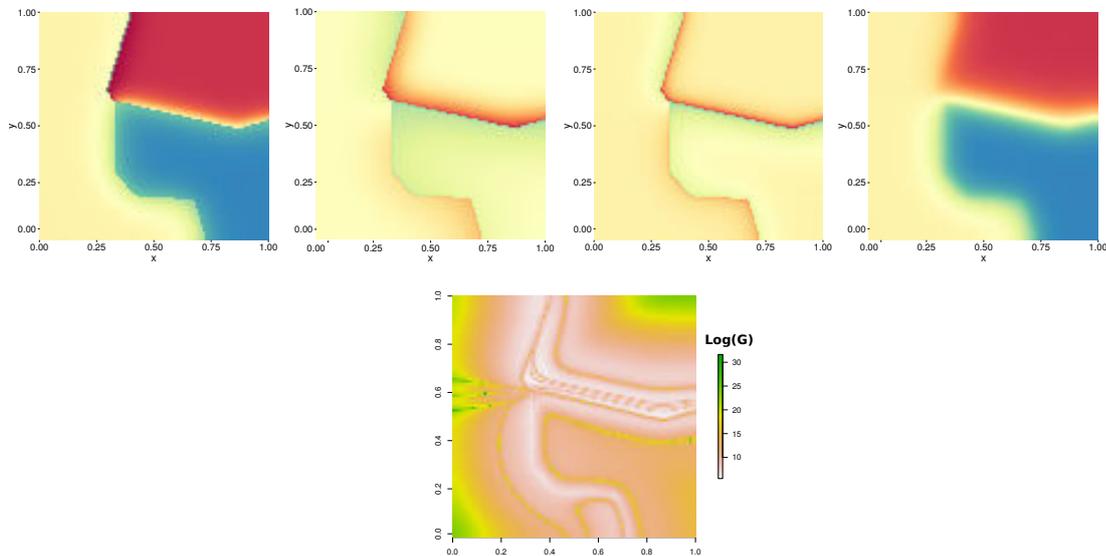


FIGURE 3.1 – Densités de population associées à 4 modèles de processus écologiques en fin de simulation (panneaux supérieurs) et distribution spatiale de la métrique de contrastes associée.

Si la figure 3.1 ne montre qu'une photographie de la métrique de contraste à un temps donné, la métrique est en réalité calculée en tout point de l'espace et du temps. Pour améliorer notre heuristique, il faudrait échantillonner les données virtuelles en des points de l'espace et du temps qui maximisent la métrique de contraste, tout en gardant en tête les contraintes de l'échantillonnage sur le terrain,

en contraignant par exemple le nombre de dates d'échantillonnage. Une première étape pour vérifier si l'ajout de cette métrique de contraste est une réelle amélioration serait de comparer les performances de classification associées à l'échantillonnage sélectionné dans l'article et un échantillonnage basé sur la métrique de contraste présentant le même effort total d'échantillonnage.

Références de la partie III

- Aben, Job, Greta Bocedi, Stephen C. F. Palmer, Petri Pellikka, Diederik Strubbe, Caspar Hallmann, Justin M. J. Travis, Luc Lens et Erik Matthysen (08/2016). “The importance of realistic dispersal models in conservation planning : application of a novel modelling platform to evaluate management scenarios in an Afrotropical biodiversity hotspot”. en. In : *Journal of Applied Ecology* 53.4, p. 1055-1065. URL : <http://onlinelibrary.wiley.com/doi/10.1111/1365-2664.12643/abstract> (cf. pp. 5, 59).
- Alema, Bas (2014). “Quantifying and simulating movement of the predator carabid beetle *Pterostichus melanarius* in arable land”. Thèse de doct. Wageningen : Wageningen University. URL : <https://www.wageningenur.nl/en/Publication-details.htm?publicationId=publication-way-343533353438> (cf. pp. 25, 31, 33, 34, 64, 66, 82, 101).
- Avgar, Tal, Rob Deardon et John M. Fryxell (02/2013). “An empirically parameterized individual based model of animal movement, perception, and memory”. In : *Ecological Modelling* 251, p. 158-172. URL : <https://www.sciencedirect.com/science/article/pii/S0304380012005686> (cf. p. 60).
- Banks, Robert B. (1994). *Growth and Diffusion Phenomena*. Sous la dir. de F. John, J. E. Marsden, L. Sirovich, M. Golubitsky et W. Jäger. T. 14. Texts in Applied Mathematics. Berlin, Heidelberg : Springer Berlin Heidelberg. URL : <http://link.springer.com/10.1007/978-3-662-03052-3> (cf. pp. 34, 65).
- Barros, Ceres, Stephen C. F. Palmer, Greta Bocedi et Justin M. J. Travis (12/2016). “Spread rates on fragmented landscapes : the interacting roles of demography, dispersal and habitat availability”. en. In : *Diversity and Distributions* 22.12, p. 1266-1275. URL : <http://onlinelibrary.wiley.com/doi/10.1111/ddi.12487/abstract> (cf. pp. 34, 65).
- Baudry, Jacques, Françoise Burel, Claudine Thenail et Didier Le Cœur (08/2000). “A holistic landscape ecological study of the interactions between farming activities and ecological patterns in Brittany, France”. In : *Landscape and Urban Planning* 50.1-3, p. 119-128. URL : <http://www.sciencedirect.com/science/article/pii/S0169204600000840> (cf. p. 64).
- Blitzer, Eleanor J., Carsten F. Dormann, Andrea Holzschuh, Alexandra-Maria Klein, Tatyana A. Rand et Teja Tschardt (01/2012). “Spillover of functionally important organisms between managed and natural habitats”. In : *Agriculture, Ecosystems & Environment* 146.1, p. 34-43. URL : <http://www.sciencedirect.com/science/article/pii/S0167880911003306> (cf. p. 83).
- Bohan, David A., Aline Boursault, David R. Brooks et Sandrine Petit (08/2011). “National-scale regulation of the weed seedbank by carabid predators”. en. In : *Journal of Applied Ecology*

- 48.4, p. 888-898. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2664.2011.02008.x/abstract> (cf. p. 67).
- Bourhis, Yoann, Sylvain Poggi, Youcef Mammeri, Anne-Marie Cortesero, Anne Le Ralec et Nicolas Parisey (09/2015). "Perception-based foraging for competing resources : Assessing pest population dynamics at the landscape scale from heterogeneous resource distribution". In : *Ecological Modelling* 312, p. 211-221. URL : <http://www.sciencedirect.com/science/article/pii/S0304380015002355> (cf. pp. 34, 65, 121).
- C F George Thomas, John M. Holland (2002). "The spatial distribution of carabid beetles in agricultural landscapes". In : *The agroecology of carabid beetles*. Intercept, p. 305-344 (cf. p. 61).
- Codling, Edward A., Michael J. Plank et Simon Benhamou (08/2008). "Random walk models in biology". en. In : *Journal of The Royal Society Interface* 5.25, p. 813-834. URL : <http://rsif.royalsocietypublishing.org/content/5/25/813> (cf. pp. 65, 100).
- Cushman, Samuel A., Kevin McGarigal et Maile C. Neel (09/2008). "Parsimony in landscape metrics : Strength, universality, and consistency". In : *Ecological Indicators* 8.5, p. 691-703. URL : <http://www.sciencedirect.com/science/article/pii/S1470160X07001306> (cf. pp. 26, 28, 64).
- Edelman, Andrew J. (02/2012). "Positive Interactions between Desert Granivores : Localized Facilitation of Harvester Ants by Kangaroo Rats". In : *PLoS ONE* 7.2. URL : <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3279350/> (cf. pp. 29, 30, 65).
- Fahrig, Lenore (01/1998). "When does fragmentation of breeding habitat affect population survival?" In : *Ecological Modelling* 105.2-3, p. 273-292. URL : <http://www.sciencedirect.com/science/article/pii/S0304380097001634> (cf. p. 59).
- Fahrig, Lenore, Jacques Baudry, Lluís Brotons, Françoise G. Burel, Thomas O. Crist, Robert J. Fuller, Clelia Sirami, Gavin M. Siriwardena et Jean-Louis Martin (02/2011). "Functional landscape heterogeneity and animal biodiversity in agricultural landscapes". en. In : *Ecology Letters* 14.2, p. 101-112. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2010.01559.x/abstract> (cf. pp. 2, 26, 59, 64).
- Fahrig, Lenore, Judith Girard, Dennis Duro, Jon Pasher, Adam Smith, Steve Javorek, Douglas King, Kathryn Freemark Lindsay, Scott Mitchell et Lutz Tischendorf (02/2015). "Farmlands with smaller crop fields have higher within-field biodiversity". In : *Agriculture, Ecosystems & Environment* 200, p. 219-234. URL : <http://www.sciencedirect.com/science/article/pii/S0167880914005192> (cf. pp. 7, 9, 26, 43, 59, 63).
- Fortin, Marie-Josée et Mark RT Dale (2005). *Spatial analysis : a guide for ecologists*. Cambridge University Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=bTT7tEiTFLoC&>

oi=fnd&pg=PA1&dq=spatial+analysis:+a+guide+for+ecologists.&ots=qSgTFacVRj&sig=gAXaISwd8Y86GLbd0tn3z67T7vA (cf. pp. 4, 60).

Friedman, Jerome H. (04/2001). "Greedy function approximation : A gradient boosting machine." In : *The Annals of Statistics* 29.5, p. 1189-1232. URL : <http://projecteuclid.org/euclid.aos/1013203451> (cf. p. 68).

Gaucherel, C. (11/2008). "Neutral models for polygonal landscapes with linear networks". In : *Ecological Modelling* 219.1-2, p. 39-48. URL : <http://www.sciencedirect.com/science/article/pii/S0304388008003803> (cf. pp. 28, 64).

Gillespie, Mark A. K. *et al.* (04/2017). "A method for the objective selection of landscape-scale study regions and sites at the national level". en. In : *Methods in Ecology and Evolution*, n/a-n/a. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12779/abstract> (cf. p. 63).

Grimm, Volker, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M. Mooij, Steven F. Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand et Donald L. DeAngelis (11/2005). "Pattern-Oriented Modeling of Agent-Based Complex Systems : Lessons from Ecology". en. In : *Science* 310.5750, p. 987-991. URL : <http://www.sciencemag.org/content/310/5750/987> (cf. pp. 8-10, 56, 60, 96, 100).

Hall, Linnea S., Paul R. Krausman et Michael L. Morrison (1997). "The Habitat Concept and a Plea for Standard Terminology". In : *Wildlife Society Bulletin (1973-2006)* 25.1, p. 173-182. URL : <http://www.jstor.org/stable/3783301> (cf. p. 65).

Hastie, Trevor J. et Robert J. Tibshirani (1990). *Generalized additive models*. T. 43. CRC Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=qa29r1Ze1coC&oi=fnd&pg=PR13&dq=Generalized+Additive+Models+hastie&ots=j32WgmzYpI&sig=10sLwH2xhu0avZiqYA-ELdUo4tg> (cf. p. 69).

Hastie, Trevor, Robert Tibshirani et Jerome Friedman (08/2009). *The Elements of Statistical Learning : Data Mining, Inference, and Prediction, Second Edition*. en. Google-Books-ID : tVIjmNS3Ob8C. Springer Science & Business Media (cf. p. 68).

Holland J. M., Begbie M. (2004). "The spatial dynamics and movement of *Pterostichus melanarius* and *P. madidus* (Carabidae) between and within arable fields in the UK". In : *International Journal of Ecology and Environmental Sciences* 30, p. 35-50 (cf. pp. 24, 61).

Holland, J. M. et M. L. Luff (06/2000). "The Effects of Agricultural Practices on Carabidae in Temperate Agroecosystems". en. In : *Integrated Pest Management Reviews* 5.2, p. 109-129. URL : <http://link.springer.com/article/10.1023/A%3A1009619309424> (cf. pp. 59, 61).

Holland, J. M., C. F. G. Thomas, T. Birkett, S. Southway et H. Oaten (12/2005). "Farm-scale spatiotemporal dynamics of predatory beetles in arable crops". en. In : *Journal of Applied*

- Ecology* 42.6, p. 1140-1152. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2664.2005.01083.x/abstract> (cf. pp. 21, 64, 67).
- Honek, Alois et Vojtech Jarosik (12/2000). "The role of crop density, seed and aphid presence in diversification of field communities of Carabidae (Coleoptera)". In : *European Journal of Entomology* 97.4, p. 517-525. URL : <http://www.eje.cz/doi/10.14411/eje.2000.080.html> (cf. p. 59).
- Huzimura, Ryoitiro et Toyoki Matsuyama (12/1999). "A Mathematical Model with a Modified Logistic Approach for Singly Peaked Population Processes". In : *Theoretical Population Biology* 56.3, p. 301-306. URL : <http://www.sciencedirect.com/science/article/pii/S0040580999914260> (cf. p. 65).
- J. Haynes, Kyle et James T. Cronin (04/2006). "Interpatch movement and edge effects : the role of behavioral responses to the landscape matrix". en. In : *Oikos* 113.1, p. 43-54. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.0030-1299.2006.13977.x/abstract> (cf. p. 83).
- Jackson, Heather Bird et Lenore Fahrig (01/2015). "Are ecologists conducting research at the optimal scale?" en. In : *Global Ecology and Biogeography* 24.1, p. 52-63. URL : <http://onlinelibrary.wiley.com/doi/10.1111/geb.12233/abstract> (cf. p. 59).
- Johnson, Jerald B. et Kristian S. Omland (02/2004). "Model selection in ecology and evolution". In : *Trends in Ecology & Evolution* 19.2, p. 101-108. URL : <http://www.sciencedirect.com/science/article/pii/S0169534703003458> (cf. pp. 35, 60, 69, 83).
- Jongejans, Eelke, Olav Skarpaas et Katriona Shea (03/2008). "Dispersal, demography and spatial population models for conservation and control management". In : *Perspectives in Plant Ecology, Evolution and Systematics*. Space matters - Novel developments in plant ecology through spatial modelling 9.3-4, p. 153-170. URL : <http://www.sciencedirect.com/science/article/pii/S1433831907000467> (cf. pp. 5, 8, 59, 61, 81).
- Jonsson, Mattias, Riccardo Bommarco, Barbara Ekbom, Henrik G. Smith, Jan Bengtsson, Berta Caballero-Lopez, Camilla Winqvist et Ola Olsson (03/2014). "Ecological production functions for biological control services in agricultural landscapes". en. In : *Methods in Ecology and Evolution* 5.3, p. 243-252. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12149/abstract> (cf. pp. 5, 59).
- Landis, Douglas A., Stephen D. Wratten et Geoff M. Gurr (2000). "Habitat Management to Conserve Natural Enemies of Arthropod Pests in Agriculture". In : *Annual Review of Entomology* 45.1, p. 175-201. URL : <http://dx.doi.org/10.1146/annurev.ento.45.1.175> (cf. pp. 5, 24, 59).

- Le Ber, Florence, Claire Lavigne, Katarzyna Adamczyk, Frédérique Angevin, Nathalie Colbach, Jean-François Mari et Hervé Monod (2009). "Neutral modelling of agricultural landscapes by tessellation methods-Application for gene flow simulation." In : *Ecological Modelling* 220, p. 3536-3545. URL : <https://hal.archives-ouvertes.fr/hal-00409081> (cf. pp. 28, 40, 64).
- Marrec, Ronan, Isabelle Badenhauer, Vincent Bretagnolle, Luca Börger, Marilyn Roncoroni, Nadine Guillon et Bertrand Gauffre (01/2015). "Crop succession and habitat preferences drive the distribution and abundance of carabid beetles in an agricultural landscape". In : *Agriculture, Ecosystems & Environment* 199, p. 282-289. URL : <http://www.sciencedirect.com/science/article/pii/S016788091400468X> (cf. pp. 7, 9, 24, 33, 67).
- Mateu, Jorge et Werner G. Müller (2012). *Spatio-temporal design : advances in efficient data acquisition*. John Wiley & Sons. URL : <https://books.google.fr/books?hl=fr&lr=&id=GWRqLPcTEnc&oi=fnd&pg=PT9&dq=Spatio-temporal+Design:+Advances+in+Efficient+Data+Acquisition&ots=DkjDUEB6XE&sig=mLzvK-A0iKU5JtwAJp3tMASAGUM> (cf. p. 60).
- McIntire, Eliot J. B. et Alex Fajardo (01/2009). "Beyond description : the active and effective way to infer processes from spatial patterns". In : *Ecology* 90.1, p. 46-56. URL : <http://www.esajournals.org/doi/abs/10.1890/07-2096.1> (cf. pp. 8, 9, 54, 56, 60).
- McIntire, Eliot J. B., Ghislain Rompré et Paul M. Severns (05/2013). "Biased correlated random walk and foray loop : which movement hypothesis drives a butterfly metapopulation?" en. In : *Oecologia* 172.1, p. 293-305. URL : <http://link.springer.com/article/10.1007/s00442-012-2475-9> (cf. pp. 10, 54, 60, 100).
- Moloney, Kirk A. et Simon A. Levin (1996). "The effects of disturbance architecture on landscape-level population dynamics". In : *Ecology*, p. 375-394. URL : <http://www.jstor.org/stable/2265616> (cf. pp. 5, 59).
- Morris, Max D. (05/1991). "Factorial Sampling Plans for Preliminary Computational Experiments". In : *Technometrics* 33.2, p. 161-174. URL : <http://www.tandfonline.com/doi/abs/10.1080/00401706.1991.10484804> (cf. pp. 70, 78, 125).
- Nathan, Ran (09/2008). "An emerging movement ecology paradigm". en. In : *Proceedings of the National Academy of Sciences* 105.49, p. 19050-19051. URL : <http://www.pnas.org/content/105/49/19050> (cf. pp. 10, 54, 60).
- Padrón, Víctor et María Cristina Trevisan (02/2006). "Environmentally induced dispersal under heterogeneous logistic growth". In : *Mathematical Biosciences* 199.2, p. 160-174. URL : <http://www.sciencedirect.com/science/article/pii/S0025556405002075> (cf. pp. 34, 65).
- Parisey, Nicolas, Yoann Bourhis, Lionel Roques, Samuel Soubeyrand, Benoît Ricci et Sylvain Poggi (12/2016). "Rearranging agricultural landscapes towards habitat quality optimisation : In silico

- application to pest regulation". In : *Ecological Complexity* 28, p. 113-122. URL : <https://www.sciencedirect.com/science/article/pii/S1476945X16300538> (cf. pp. 5, 34, 59, 101).
- Platt, J. R. (10/1964). "Strong Inference : Certain systematic methods of scientific thinking may produce much more rapid progress than others". eng. In : *Science (New York, N.Y.)* 146.3642, p. 347-353 (cf. pp. 56, 60, 83).
- Roques, Lionel, Marie-Anne Auger-Rozenberg et Alain Roques (11/2008). "Modelling the impact of an invasive insect via reaction-diffusion". In : *Mathematical Biosciences* 216.1, p. 47-55. URL : <http://www.sciencedirect.com/science/article/pii/S0025556408001235> (cf. pp. 33, 65).
- Roques, Lionel et M. D. Chekroun (03/2010). "Does reaction-diffusion support the duality of fragmentation effect?" In : *Ecological Complexity* 7.1, p. 100-106. URL : <http://www.sciencedirect.com/science/article/pii/S1476945X09000646> (cf. p. 65).
- Savage, David et Michael Renton (01/2014). "Requirements, design and implementation of a general model of biological invasion". In : *Ecological Modelling* 272, p. 394-409. URL : <http://www.sciencedirect.com/science/article/pii/S0304380013004638> (cf. pp. 5, 59).
- Schröder, B. et R. Seppelt (12/2006). "Analysis of pattern-process interactions based on landscape models—Overview, general concepts, and methodological issues". In : *Ecological Modelling. Pattern and Processes of Dynamic Mosaic Landscapes – Modelling, Simulation, and Implications* 199.4, p. 505-516. URL : <http://www.sciencedirect.com/science/article/pii/S0304380006002936> (cf. pp. 3, 4, 59-61, 96).
- Schweiger, Andreas H., Severin D. H. Irl, Manuel J. Steinbauer, Jürgen Dengler et Carl Beierkuhnlein (04/2016). "Optimizing sampling approaches along ecological gradients". en. In : *Methods in Ecology and Evolution* 7.4, p. 463-471. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12495/abstract> (cf. pp. 10, 60, 101).
- Shigesada, Nanako et Kohkichi Kawasaki (02/1997). *Biological Invasions : Theory and Practice*. en. Oxford University Press, UK (cf. pp. 33, 65).
- Tilman, David et Peter M. Kareiva (1997). *Spatial ecology : the role of space in population dynamics and interspecific interactions*. T. 30. Princeton University Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=QdqC8PvWB7AC&oi=fnd&pg=PR7&dq=Spatial+ecology:+the+role+of+space+in+population+dynamics+and+interspecific+interactions.&ots=bKYtHjZj&sig=4VpUdp00mt23W37oQZ6aRi7BJAY> (cf. pp. 8, 60).
- Tuda, Midori (03/2007). "Understanding mechanism of spatial ecological phenomena : a preface to the special feature on "Spatial statistics"". en. In : *Ecological Research* 22.2, p. 183-184. URL : <http://link.springer.com/article/10.1007/s11284-006-0328-x> (cf. pp. 8, 60).

- Turner, Monica G. (2005). "LANDSCAPE ECOLOGY : What Is the State of the Science?" In : *Annual Review of Ecology, Evolution, and Systematics* 36.1, p. 319-344. URL : <http://dx.doi.org/10.1146/annurev.ecolsys.36.102003.152614> (cf. p. 59).
- Vinatier, Fabrice, Philippe Tixier, Pierre-François Duyck et Françoise Lescourret (2011). "Factors and mechanisms explaining spatial heterogeneity : a review of methods for insect populations". en. In : *Methods in Ecology and Evolution* 2.1, p. 11-22. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2010.00059.x/abstract> (cf. pp. 9, 81, 100).
- Wood, Simon N. (02/2003). "Thin plate regression splines". en. In : *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 65.1, p. 95-114. URL : <http://onlinelibrary.wiley.com/doi/10.1111/1467-9868.00374/abstract> (cf. p. 69).
- Zurell, Damaris *et al.* (04/2010). "The virtual ecologist approach : simulating data and observers". en. In : *Oikos* 119.4, p. 622-635. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0706.2009.18284.x/abstract> (cf. pp. 10, 54, 60).

Quatrième partie

Discussion générale et conclusion

4.1	Bilan des travaux	96
4.2	Lien entre processus écologiques et services écosytémiques	101
4.3	Conclusion	104

4.1 Bilan des travaux

Le processus d'acquisition de connaissances sur la relation entre processus écologiques et motifs de population est une démarche itérative impliquant des allers retours entre observations sur le terrain et modélisation des processus (Holling et Allen 2002; Peck 2004). À la base de cette démarche, des études empiriques d'observation des motifs de population mènent à la formulation d'hypothèses sur les processus ayant mené à ces observations. La diversité des études empiriques en termes d'échelles spatiales et temporelles d'observation, de variables paysagères extraites, de méthodes pour l'échantillonnage biologique et de traitement statistique des résultats génèrent une diversité d'hypothèses sur les processus écologiques sous-jacents parfois en contradiction les unes avec les autres.

Ces différentes hypothèses doivent ensuite être traduites en modèles de dynamiques de population et leur trajectoire spatio-temporelle étudiée. La sélection de l'hypothèse écologique la plus plausible se fait alors sur la base de la comparaison entre les motifs de population générés et de nouvelles observations, aux échelles temporelles et spatiales jugées pertinentes par l'analyse des motifs générés par la simulation. Si l'adéquation entre motifs simulés et observés n'est pas satisfaisante, les hypothèses initiales peuvent être affinées ou de nouvelles hypothèses peuvent émerger (Grimm *et al.* 2005).

Une fois qu'une hypothèse a été jugée suffisamment pertinente pour expliquer les motifs de population observés, le modèle de dynamique de population sélectionné peut être utilisé dans le cadre de la modélisation directe pour l'exploration de scénarios paysagers dans une optique de gestion de l'espèce (Schröder et Seppelt 2006).

Le présent travail de thèse s'inscrit dans cette démarche d'acquisition de connaissance sur le lien entre processus et motifs en se focalisant sur la partie centrale liée à la modélisation des processus écologiques sur un espace spatialement explicite.

Dans un premier temps (chapitre II), nous avons cherché à voir si cette démarche pouvait être court-circuitée pour déterminer si la somme des connaissances disponibles sur l'organisme d'intérêt permettait directement d'amener à des conclusions sur l'effet d'une variable paysagère. En confrontant la diversité des hypothèses écologiques à un gradient de la variable paysagère et en comparant les motifs générés, il est alors possible de déterminer si les processus conjecturés amènent à des réponses similaires. Si c'est le cas, et à la condition que les valeurs des paramètres de dynamiques de population soient dans une gamme réaliste, il n'est alors pas nécessaire d'améliorer la connaissance sur les processus écologiques avant d'explorer les scénarios paysagers dans une optique de gestion de l'espèce. Cette réponse de la dynamique de population à la variable paysagère, unique en dépit de la diversité des

processus sous-jacents, peut-être utilisée directement dans le processus de réflexion sur l'aménagement du paysage.

Lorsque la connaissance entre processus et motifs doit être améliorée, la sélection des hypothèses les plus pertinentes doit s'appuyer sur la comparaison entre les motifs simulés et les motifs observés sur le terrain. Dans le chapitre III, nous avons mis en place une heuristique pour la sélection de stratégies d'échantillonnage permettant d'extraire des motifs de population contenant suffisamment d'information pour sélectionner les processus écologiques régissant les dynamiques de population de l'organisme d'intérêt.

Les efforts d'échantillonnage habituellement consentis dans des études en écologie du paysage (voir *e.g.* Bertrand (2015) et Puech (2014) pour des études effectuées sur la Zone Atelier Armorique) ont été explorés pour étudier leur potentiel à fournir des données biologiques permettant une inférence robuste des processus écologiques.

Nous avons montré que la manière de distribuer l'effort d'échantillonnage entre ses composantes spatiale et temporelle était un levier majeur sur la capacité à discriminer les hypothèses écologiques formulées *a priori*. Pour un effort donné, le choix de favoriser l'une ou l'autre des dimensions peut amener à une variabilité élevée des performances de classification. Dans le même temps, pour un même effort, des protocoles d'échantillonnage différant dans leurs dimensions spatiales et temporelles peuvent amener à des performances similaires, et le choix de l'échantillonnage peut alors s'effectuer sur la base des contraintes logistiques liées à la campagne de terrain (*e.g.* nombre de pièges à disposition, nombre de visites possibles sur la zone d'étude, nombre de paysages où l'échantillonnage est possible).

Le modèle additif généralisé (cf. Partie 2, §2.2.4), qui ajuste la relation entre les performances de classification et les dimensions de l'échantillonnage, a permis de mettre en évidence la nature non-linéaire de cette relation. Ce modèle montre également que le choix de la dimension à privilégier pour améliorer les performances de classification dépend de la position dans l'espace des dimensions d'échantillonnage. En somme, ce modèle permet d'atteindre un optimum par la recherche de maxima locaux sujets à des contraintes d'échantillonnage définies *a priori*.

Le processus d'acquisition de connaissances sur les processus écologiques, dans lequel s'inscrivent ces travaux de thèse, est résolument pluridisciplinaire car il consiste en une succession d'étapes impliquant de nombreux champs de l'écologie du paysage et des disciplines associées (cf. figure 1.4) : les statistiques spatiales et spatio-temporelles pour l'analyse des motifs de population, la comparaison des motifs simulés et observés et l'optimisation des stratégies d'échantillonnage ; l'analyse et la modélisation de paysages pour la génération des instances de paysages virtuels et l'exploration de scénarios de structures paysagères ; la modélisation de dynamiques

de population pour la génération de données biologiques virtuelles ; la classification statistique pour le regroupement des modèles de dynamiques de population L'originalité de cette thèse provient de l'association de ces différents champs disciplinaires pour en bâtir un raisonnement méthodologique visant à améliorer les connaissances sur les relations entre processus écologiques et les motifs de populations. Cependant, si elle est l'essence et l'intérêt de ces travaux, la pluridisciplinarité s'accompagne souvent, du fait de nos limites cognitives (tout au moins, des miennes), d'une impossibilité d'être expert dans toutes les disciplines abordées. Ainsi les méthodes proposées à chaque étape de notre raisonnement sont elles perfectibles, et certains choix méthodologiques critiquables par des esprits plus éclairés. Nous proposons ici quelques pistes pour l'amélioration de certaines étapes de nos travaux.

4.1.1 Caractérisation des motifs de population

La quantité de données générées par la simulation de multiples modèles de dynamiques de populations, à des résolutions élevées en temps et en espace, nous a contraints à utiliser des métriques de population agrégées en temps et en espace pour être en mesure de comparer les sorties des différents modèles de processus. Comme nous l'avons discuté dans le chapitre III, l'utilisation de statistiques résumées en temps et en espace amène à une perte d'informations sur les dynamiques fines des populations, et a donc tendance à réduire les contrastes entre les motifs de populations associés à chaque processus.

Dans le cadre du chapitre II, si les métriques agrégées ont un sens en termes de services écosystémiques, il est pertinent de travailler avec ces métriques, car elles déterminent directement le lien entre le besoin de connaissances sur les processus écologiques et l'échelle de gestion de l'organisme d'intérêt. Nous discutons cela plus en détail dans la section 4.2.

Dans le cadre du chapitre III, en revanche, où la discrimination entre les modèles supportant les hypothèses écologiques est basée sur les distances entre les motifs de population, cette perte d'information rend plus difficile la discrimination entre les processus sous-jacents. Comme nous l'évoquons dans la section 3, il est possible qu'utiliser directement les données brutes (*i.e.* densités ou abondances en $\{x, y, t\}$) en lieu et place de statistiques agrégées permette d'améliorer la sélection des processus écologiques pour des efforts d'échantillonnage équivalents. Comparer les performances de classification selon que les données d'entrée sont les données brutes ou des statistiques résumées en temps et en espace serait une vérification importante à effectuer dans le cadre de notre travail. Utiliser les données brutes pose cependant le problème du sur-apprentissage statistique - le nombre de variables devenant trop largement supérieur au nombre de modèles de processus à classifier - qui devra

préalablement être résolu.

4.1.2 Incertitude et sensibilité

Dans tout travail de modélisation, il est une nécessité d'identifier les sources d'incertitude liées à notre manque de connaissance sur le système modélisé, et de quantifier *a priori* l'impact de ces incertitudes sur les sorties du modèle (Faivre *et al.* 2013). En d'autres termes, il convient d'évaluer l'impact sur les sorties Y des composantes du modèle $Y = f(x, \theta)$, où x est le vecteur des variables d'entrée, incluant la population initiale et les variables paysagères, θ le vecteur des paramètres et f l'équation représentant le processus à modéliser.

Selon cette définition, les travaux présentés dans le chapitre II peuvent être assimilés à une analyse d'incertitude liée à la formulation des processus (la fonction f de l'équation précédente) et à la variable paysagère sur les descripteurs de population. La notion de sensibilité à la formulation des processus est ensuite approchée par la classification : les descripteurs de population des modèles appartenant à une même classe sont considérés comme non sensibles à la formalisation de f .

Pour rendre complète cette étude, il conviendrait d'évaluer également l'impact des valeurs des paramètres des modèles de processus et des conditions initiales de population sur les descripteurs de population. Une perspective intéressante serait alors de réaliser une analyse de sensibilité des paramètres de dynamique de population, des conditions initiales et des variables paysagères sur les descripteurs de population, pour chaque modèle de processus. Cela permettrait de déterminer, pour chaque hypothèse sur les processus écologiques, si la structure du paysage est un levier réel sur les descripteurs de population. En effet, si pour un modèle de processus donné, l'effet de la variable paysagère est négligeable, il n'apparaîtra pas comme prioritaire de jouer sur la structure du paysage pour favoriser cette espèce. A l'inverse, si la variable paysagère apparaît comme la variable la plus sensible, améliorer la connaissance sur les processus écologiques, *i.e.* mieux calibrer les paramètres de population ne sera pas prioritaire dans une optique de gestion de l'espèce.

Malheureusement, le nombre élevé de variables et de paramètres rendent difficile l'analyse de sensibilité des modèles couplés de paysage/dynamique de population. Effectuer des simulations spatialement explicites des modèles de dynamique de population, déjà individuellement coûteuses en ressources de calcul, sur l'ensemble de l'espace de variation des paramètres biologiques et des variables paysagères, en y ajoutant la réplication des paysages pour prendre en compte la variabilité des paysages pour une même valeur (Papaix *et al.* 2014), demeure un défi technique à surmonter. La méta-modélisation, *i.e.* la formulation de modèles simplifiés sur la base de l'interpolation globale des comportements locaux du modèle complexe, peut être

une perspective intéressante, mais elle ne garantit pas que l'ensemble des comportements du modèle en certains points de l'espace des paramètres soient correctement capturés (Faivre *et al.* 2013).

4.1.3 Formalisme des modèles de dynamique des populations

La population est un ensemble d'individus dont les comportements présentent une variabilité due à l'interaction complexe de facteurs exogènes biotiques ou abiotiques (*e.g.* conditions (micro-)climatiques, type d'occupation du sol, interactions avec d'autres individus) et endogènes (*e.g.* état de satiété, variabilité phénotypique) (Kromp 1999). La somme de ces comportements individuels, l'occurrence d'événements rares due à la stochasticité de ces comportements, peuvent avoir des répercussions sur les motifs de population qui ne peuvent être capturés par un formalisme qui traite des dynamiques de l'organisme à l'échelle populationnelle (Peck 2004). Le formalisme de réaction-diffusion approche la somme des comportements individuels par des équations à l'échelle populationnelle (Codling *et al.* 2008; Vinatier *et al.* 2011). En ce sens, il est susceptible de décrire les processus écologiques d'une manière trop simpliste pour faire émerger des motifs de population que pourraient faire apparaître des modèles individus centrés (IBM pour Individual-Based Models, DeAngelis et Grimm (2014)), capables de prendre en compte la variabilité des comportements (inter-)individuels. Parce qu'il permet d'intégrer une plus grande complexité, ce formalisme a parfois été utilisé pour inférer les processus écologiques à partir de l'étude des motifs, par le biais de l'approche de *pattern-oriented modelling* (Grimm *et al.* 2005; McIntire, Rompré *et al.* 2013).

Cependant, si se placer à l'échelle individuelle peut permettre de mieux répondre à la problématique de l'inférence des processus à partir des motifs de population, le formalisme IBM est coûteux, tant en termes de connaissances initiales sur les processus écologiques - et donc dans le nombre de paramètres à tester - qu'en termes de temps de calcul. À l'heure actuelle, adapter notre heuristique pour utiliser ce formalisme est impossible du fait du nombre élevé de simulations impliquées, surtout en considérant le besoin de quantifier les incertitudes décrit dans la section précédente 4.1.2.

4.1.4 Prise en compte du biais d'observation lié au mode d'échantillonnage

Dans le chapitre III, nous avons présenté un cadre méthodologique pour déterminer une stratégie d'échantillonnage permettant de sélectionner un processus écologique parmi un ensemble d'hypothèses a priori. La continuité de ce travail est

d'acquérir les données biologiques par la stratégie sélectionnée, puis de comparer ces données observées aux données simulées pour en déduire quel processus représente au mieux les dynamiques de l'espèce d'intérêt sur le terrain.

Cependant, les données échantillonnées sur le terrain ne sont que rarement directement comparables aux données générées par les modèles. En effet, ce qui est observé sur le terrain n'est pas *per se* la grandeur d'intérêt (*e.g.* une densité d'individus par unité de surface), mais une approximation de cette grandeur. Les données peuvent être bruitées, du fait de la composante stochastique du processus d'observation, et/ou le mode d'acquisition des données introduit un biais systématique entre l'observation et la réalité (Schweiger *et al.* 2016). Pour pouvoir comparer les données observées et les données simulées, il est donc nécessaire d'appliquer une correction aux observations.

Un ensemble de méthodes, regroupées sous le terme générique de *modèles d'observation*, consiste à appliquer une fonction sur le modèle de processus afin de reproduire virtuellement le processus d'observation (Clark 2005 ; Soubeyrand et Roques 2014). Des modèles d'observation ont ainsi été utilisés pour découpler les composantes d'activité et de densité induite des carabes lors de l'échantillonnage par les pots Barber, allant d'une simple correction linéaire en fonction de la température et de la densité du couvert (Honek 1988), à des modèles mécanistico-statistiques plus complexes (Alema 2014). Intégrer un modèle d'observation à notre cadre méthodologique permettrait de prendre en compte les biais d'observation liés à l'échantillonnage des carabes épigés par pots Barber (Parisey *et al.* 2016), mais également de l'adapter à d'autres types de mode d'acquisition de données biologiques (*e.g.* prélèvement d'insectes sur les parties aériennes des plantes par filets fauchoirs, relevé aviaires par points d'écoutes, *etc.*).

4.2 Lien entre processus écologiques et services écosystémiques

4.2.1 Description des dynamiques biologiques selon des critères liés aux services écosystémiques

Dans le chapitre II, la notion de « réponse similaire » à une variable paysagère a été abordée par le prisme d'un problème de classification. Les modèles de processus ont été classés sur la base des distances entre les descripteurs de population associés à chaque modèle. Les modèles regroupés dans une même classe sont ceux qui ont de faibles distances entre eux, relativement à la distance avec les modèles des autres classes. En d'autres termes, la discrimination des modèles est basée sur des critères

relatifs qui n'ont pas forcément de sens fonctionnel. Dans une problématique de gestion d'espèce, où les descripteurs de population peuvent avoir un sens en termes de services écosystémiques rendus par l'organisme d'intérêt, il serait pertinent de déterminer si la connaissance sur les processus écologiques se doit d'être améliorée sur la base de ces critères fonctionnels. Au lieu d'une classification non supervisée, l'utilisateur peut, sur la base de dire d'expert sur le lien entre les descripteurs de population et le service écosystémique d'intérêt, déterminer des valeurs seuils de distance entre les descripteurs en deçà desquelles les modèles sont considérés comme appartenant à une même classe (Yang 2006).

D'une manière plus générale, le besoin ou non d'améliorer les connaissances sur les processus écologiques peut se faire sur la base de critères plus fonctionnels que des descripteurs agrégés sur l'ensemble de la saison et à l'échelle du paysage. Dans notre cas, où nous nous intéressons à un organisme pour son potentiel de régulation d'un ravageur, il serait pertinent d'extraire des descripteurs de population du prédateur à des moments clés de la dynamique du ravageur, *e.g.* la densité de prédateurs à l'échelle du paysage avant le pic de démographie du ravageur (Bertrand *et al.* 2016). De la même manière que pour la temporalité, les descripteurs de population du prédateur peuvent être extraits à des endroits clés du paysage en lien avec la distribution spatiale du ravageur. En somme, ce sont les attendus en termes de services écosystémiques qui déterminent le besoin ou non d'acquérir des connaissances supplémentaires sur les processus écologiques : si les motifs de population du prédateur sont similaires à des moments et des endroits clés de la dynamique du ravageur, il n'est pas nécessaire d'améliorer les connaissances même si les dynamiques sont différentes à l'échelle du paysage et sur l'ensemble de la saison.

Dans ce contexte, notre heuristique peut s'appliquer si les connaissances concernant les processus écologiques du ravageur et de leur réponse au prédateur sont elles aussi méconnues. En partant de la même manière d'hypothèses en compétition sur le système proie-prédateur, il est possible de formuler un ensemble de modèles couplés utilisant les équations de Lotka-Volterra (Holmes *et al.* 1994). Ici aussi, ces modèles seraient simulés sur un gradient paysager, et le besoin d'acquérir ou non des connaissances sur le système proie-prédateur serait basé sur la comparaison des motifs de population du ravageur et non sur ceux du prédateur.

4.2.2 Une ouverture vers un système multi-espèces

L'ensemble des travaux présentés dans cette thèse se sont focalisés sur la démarche d'acquisition de connaissances sur les processus écologiques d'une espèce unique, justifiée par un intérêt en termes de service écosystémique de régulation des ravageurs.

L'agro-écosystème, cependant, est un écosystème bien plus complexe, où de multiples espèces coexistent, potentiellement interagissent et fournissent une diversité de services écosystémiques (Bommarco *et al.* 2013). Du fait de la multiplicité des espèces qui les portent, et de leurs potentielles interactions, les services écosystémiques peuvent être en synergie ou en antagonisme (Bennett *et al.* 2009 ; Vialatte *et al.* 2016), et même au sein d'un unique type de service, la diversité des espèces peut avoir un effet négatif sur la qualité du service. Par exemple, dans une étude mettant en lien la complexité du paysage (*i.e.* la quantité d'habitats semi-naturels) et le contrôle des ravageurs par les ennemis naturels, Martin *et al.* (2013) ont montré que l'augmentation de la complexité du paysage, en favorisant la biodiversité et les interactions entre les ennemis naturels (ici la prédation) peut avoir un effet négatif sur la qualité du service de régulation des ravageurs. Cet antagonisme potentiel pose la question de compromis entre services écosystémiques, qui s'ajoute à la difficulté de la prise de décision en termes de gestion du paysage.

Aborder cette question de la recherche de paysages favorisant une diversité de services écosystémiques implique de passer de l'étude de la dynamique des populations à celle de la dynamique des communautés. Idéalement, cette recherche de paysages optimaux pourrait se faire par le biais d'une modélisation des dynamiques de population des espèces d'intérêt (*i.e.* celles fournissant des services écosystémiques et celles interagissant avec ces espèces) en intégrant les interactions entre ces espèces sur des paysages virtuels présentant une diversité de composition et/ou de configuration. S'il représente la seule manière d'approcher une vision intégrée du fonctionnement de l'agroécosystème et des services qu'il peut fournir (Bommarco *et al.* 2013), ce changement d'échelle biologique implique une augmentation drastique de la complexité de son étude. Cette démarche à long terme implique l'amélioration des connaissances sur les processus écologiques des espèces prises individuellement (l'objet de cette thèse), mais également de la nature et de l'intensité des interactions entre les espèces.

Une première étape pour aborder cette problématique est l'étude de l'effet de la structure du paysage sur un peuplement, *i.e.* sur plusieurs espèces, en s'affranchissant dans un premier temps des interactions interspécifiques. Dans cette optique, les travaux effectués dans le cadre du chapitre II peuvent constituer une première approche à cette perspective. Au lieu de considérer les modèles de dynamiques de populations comme des hypothèses en compétition concernant les processus écologiques d'une unique espèce, ces modèles peuvent représenter les dynamiques de différents organismes. Leur simulation spatialement explicite permet de générer des données de population dans l'espace et le temps, qui peuvent être comme précédemment résumées par des statistiques spatiales, temporelles et/ou spatio-temporelles, représentatives de la qualité des services écosystémiques fournis. Les dynamiques

peuvent également être caractérisées à l'échelle du peuplement via des indices de biodiversité (*e.g.* indice de Shannon, Piélou, (Bandeira *et al.* 2013)). La recherche de paysages favorisant une diversité de services écosystémiques peut alors se faire via l'exploration de scénarios de paysages en répondant à un problème d'optimisation multicritères (*e.g.* recherche du front de Pareto (Othmani 1998 ; Teillard 2012), les critères étant les métriques caractérisant les dynamiques à l'échelle des populations et à celle du peuplement.

4.3 Conclusion

Ce travail de thèse s'inscrit dans une démarche d'acquisition de connaissances sur les processus écologiques, avec pour objectif d'alimenter des modèles de dynamique des populations qui permettront *in fine* d'explorer des scénarios paysagers dans une optique de gestion d'espèces. Le contexte sociétal dans lequel s'inscrit cette démarche scientifique est celui de l'agro-écologie, qui s'emploie à utiliser les dynamiques du vivant pour remplir des fonctions écosystémiques catalysant les fonctions de production du paysage agricole. Replacer ce travail dans son contexte sociétal peut générer un sentiment d'impuissance, tant les obstacles sont nombreux pour comprendre et *in fine*, maîtriser les dynamiques de l'agro-écosystème pour maintenir les fonctions de production tout en respectant la biodiversité. Cette démarche implique en effet une succession d'étapes, présentant chacune leurs verrous scientifiques qui devront à terme être débloqués.

En premier lieu, il est nécessaire d'améliorer nos connaissances sur le lien entre les traits de vie et les éléments constituant le paysage, à l'échelle non seulement de la population (l'objet, toujours perfectible, de ce travail de thèse) mais à l'échelle de la communauté. Ce changement d'échelle augmente drastiquement la complexité du problème, car quantifier l'effet des interactions interspécifiques sur les populations en les découplant de celui des variables environnementales demeure un défi majeur de l'écologie actuelle.

Par ailleurs, la « valeur » écosystémique d'une espèce n'est pas forcément directement une fonction de son abondance et peut être difficile à quantifier (*e.g.* la présence d'un ennemi naturel supposé dans un paysage ne garantit pas l'efficacité de sa prédation, encore moins son impact positif sur les rendements agricoles). Ensuite, la recherche de paysages favorisant un bouquet de services écosystémiques par le biais de la modélisation présente une complexité méthodologique (paramétrisation de multiples modèles mécanistes en interaction, choix des variables paysagères, recherche d'optima paysagers, analyse d'incertitude et de sensibilité) et soulève la question des compromis entre services écosystémiques.

Une fois que des paysages sont jugés optimaux au vu des objectifs d'une étude, se

pose également la question de leur faisabilité en adéquation avec les contraintes de production et de viabilité économique à l'échelle du paysage et à celle des agriculteurs, qui sont les principaux acteurs du paysage agricole. Intégrer ces contraintes socio-économiques dans les modèles de paysage est possible, mais implique une augmentation drastique de la dimensionnalité du problème d'optimisation. Enfin, l'applicabilité de la démarche d'agro-écologie à l'échelle du paysage implique de prendre en compte la distorsion qu'il existe entre son échelle de gestion (*i.e.* l'exploitation agricole, un fragment du paysage) et la (les) échelle(s) des processus biologiques. La mise en place de paysages écologiquement favorables implique *a fortiori* une coopération entre les agriculteurs pour une gestion concertée du paysage dans ses dimensions spatiale et temporelle.

La mise en place d'une agriculture durable et intégrative de l'ensemble des processus écologiques, sociétaux et économiques est complexe et implique de nombreux acteurs aux objectifs parfois divergents. Le processus d'acquisition des connaissances nécessaires pour utiliser le paysage comme un levier de l'amélioration des services écosystémiques s'inscrit dans la durée, et sa temporalité semble incompatible avec l'urgence de ces objectifs. Si le rôle de la science est à mon sens d'apporter son expertise à la société civile pour favoriser son mieux-être, il me semble que la société civile peut en retour être un moteur pour accompagner ces changements. Une agriculture basée sur des agro-écosystèmes résilients et écologiquement fonctionnels ne pourra - c'est mon avis personnel - répondre à la demande future sans remise en question de nos modes de consommation et de nos comportements alimentaires qui sont, *in fine*, l'ultime levier sur la demande de production agricole.

Références citées dans la thèse

- Aben, Job, Greta Bocedi, Stephen C. F. Palmer, Petri Pellikka, Diederik Strubbe, Caspar Hallmann, Justin M. J. Travis, Luc Lens et Erik Matthysen (08/2016). “The importance of realistic dispersal models in conservation planning : application of a novel modelling platform to evaluate management scenarios in an Afrotropical biodiversity hotspot”. en. In : *Journal of Applied Ecology* 53.4, p. 1055-1065. URL : <http://onlinelibrary.wiley.com/doi/10.1111/1365-2664.12643/abstract> (cf. pp. 5, 59).
- Alema, Bas (2014). “Quantifying and simulating movement of the predator carabid beetle *Pterostichus melanarius* in arable land”. Thèse de doct. Wageningen : Wageningen University. URL : <https://www.wageningenur.nl/en/Publication-details.htm?publicationId=publication-way-343533353438> (cf. pp. 25, 31, 33, 34, 64, 66, 82, 101).
- Allema, A. B., W. van der Werf, J. C. J. Groot, L. Hemerik, G. Gort, W. a. H. Rossing et J. C. van Lenteren (02/2015). “Quantification of motility of carabid beetles in farmland”. ENG. In : *Bulletin of Entomological Research*, p. 1-11 (cf. p. 121).
- Allema, Bas, Wopke van der Werf, Joop C. van Lenteren, Lia Hemerik et Walter A. H. Rossing (12/2014). “Movement Behaviour of the Carabid Beetle *Pterostichus melanarius* in Crops and at a Habitat Interface Explains Patterns of Population Redistribution in the Field”. In : *PLoS ONE* 9.12, e115751. URL : <http://dx.doi.org/10.1371/journal.pone.0115751> (cf. p. 9).
- Avgar, Tal, Rob Deardon et John M. Fryxell (02/2013). “An empirically parameterized individual based model of animal movement, perception, and memory”. In : *Ecological Modelling* 251, p. 158-172. URL : <https://www.sciencedirect.com/science/article/pii/S0304380012005686> (cf. p. 60).
- Aviron, S., F. Burel, J. Baudry et N. Schermann (06/2005). “Carabid assemblages in agricultural landscapes : impacts of habitat features, landscape context at different spatial scales and farming intensity”. In : *Agriculture, Ecosystems & Environment*. Agri-Environmental Schemes as Landscape Experiments Agri-Environmental Schemes as Landscape Experiments 108.3, p. 205-217. URL : <http://www.sciencedirect.com/science/article/pii/S0167880905000514> (cf. p. 7).
- Baars, M. A. (12/1979). “Patterns of movement of radioactive carabid beetles”. en. In : *Oecologia* 44.1, p. 125-140. URL : <https://link.springer.com/article/10.1007/BF00346411> (cf. p. 24).
- Baddeley, Adrian (12/2010). *Analysing spatial point patterns in 'R'*. Australia. URL : <http://www.csiro.au/files/files/pn0y.pdf> (cf. p. 30).
- Bandeira, Benjamin, Jean-Louis Jamet, Dominique Jamet et Jean-Marc Ginoux (06/2013). “Mathematical convergences of biodiversity indices”. In : *Ecological Indicators* 29.Supplement C, p. 522-528. URL : <http://www.sciencedirect.com/science/article/pii/S1470160X13000605> (cf. p. 104).
- Banks, Robert B. (1994). *Growth and Diffusion Phenomena*. Sous la dir. de F. John, J. E. Marsden, L. Sirovich, M. Golubitsky et W. Jäger. T. 14. Texts in Applied Mathematics. Berlin, Heidelberg : Springer Berlin Heidelberg. URL : <http://link.springer.com/10.1007/978-3-662-03052-3> (cf. pp. 34, 65).

- Barros, Ceres, Stephen C. F. Palmer, Greta Bocedi et Justin M. J. Travis (12/2016). “Spread rates on fragmented landscapes : the interacting roles of demography, dispersal and habitat availability”. en. In : *Diversity and Distributions* 22.12, p. 1266-1275. URL : <http://onlinelibrary.wiley.com/doi/10.1111/ddi.12487/abstract> (cf. pp. 34, 65).
- Baudry, Jacques, Françoise Burel, Claudine Thenail et Didier Le Cœur (08/2000). “A holistic landscape ecological study of the interactions between farming activities and ecological patterns in Brittany, France”. In : *Landscape and Urban Planning* 50.1–3, p. 119-128. URL : <http://www.sciencedirect.com/science/article/pii/S0169204600000840> (cf. p. 64).
- Bennett, Elena M., Garry D. Peterson et Line J. Gordon (12/2009). “Understanding relationships among multiple ecosystem services”. en. In : *Ecology Letters* 12.12, p. 1394-1404. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2009.01387.x/abstract> (cf. p. 103).
- Bertrand, Colette (2015). “L’hétérogénéité spatiale et temporelle des paysages agricoles influence les auxiliaires généralistes des cultures et le potentiel de contrôle biologique des ravageurs”. Thèse de doct. Université de Rennes 1. URL : <http://www.theses.fr/194614131> (cf. p. 97).
- Bertrand, Colette, Jacques Baudry et Françoise Burel (12/2016). “Seasonal variation in the effect of landscape structure on ground-dwelling arthropods and biological control potential”. In : *Basic and Applied Ecology* 17.8, p. 678-687. URL : <http://www.sciencedirect.com/science/article/pii/S1439179116301013> (cf. p. 102).
- Blitzer, Eleanor J., Carsten F. Dormann, Andrea Holzschuh, Alexandra-Maria Klein, Tatyana A. Rand et Teja Tscharntke (01/2012). “Spillover of functionally important organisms between managed and natural habitats”. In : *Agriculture, Ecosystems & Environment* 146.1, p. 34-43. URL : <http://www.sciencedirect.com/science/article/pii/S0167880911003306> (cf. p. 83).
- Bohan, David A., Aline Boursault, David R. Brooks et Sandrine Petit (08/2011). “National-scale regulation of the weed seedbank by carabid predators”. en. In : *Journal of Applied Ecology* 48.4, p. 888-898. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2664.2011.02008.x/abstract> (cf. p. 67).
- Bommarco, Riccardo, David Kleijn et Simon G. Potts (04/2013). “Ecological intensification : harnessing ecosystem services for food security”. In : *Trends in Ecology & Evolution* 28.4, p. 230-238. URL : <http://www.sciencedirect.com/science/article/pii/S016953471200273X> (cf. pp. 20, 24, 103).
- Bourhis, Yoann, Sylvain Poggi, Youcef Mammeri, Anne-Marie Cortesero, Anne Le Ralec et Nicolas Parisey (09/2015). “Perception-based foraging for competing resources : Assessing pest population dynamics at the landscape scale from heterogeneous resource distribution”. In : *Ecological Modelling* 312, p. 211-221. URL : <http://www.sciencedirect.com/science/article/pii/S0304380015002355> (cf. pp. 34, 65, 121).
- Bourhis, Yoann, Sylvain Poggi, Youcef Mammeri, Ronan Le Cointe, Anne Marie Cortesero et Nicolas Parisey (03/2017). “Foraging as the landscape grip for population dynamics—A mechanistic model applied to crop protection”. In : *Ecological Modelling* 354, p. 26-36 (cf. p. 5).
- Burel, Françoise et Jacques Baudry (11/1999). *Ecologie du paysage. Concepts, méthodes et applications*. Français. Paris : Tec & Doc Lavoisier (cf. p. 2).

- C F George Thomas, John M. Holland (2002). “The spatial distribution of carabid beetles in agricultural landscapes”. In : *The agroecology of carabid beetles*. Intercept, p. 305-344 (cf. p. 61).
- CÁrcamo, Hector A. et John R. Spence (06/1994). “Crop Type Effects on the Activity and Distribution of Ground Beetles (Coleoptera : Carabidae)”. In : *Environmental Entomology* 23.3, p. 684-692. URL : <https://academic.oup.com/ee/article/23/3/684/417992> (cf. p. 24).
- Chaplin-Kramer, Rebecca, Megan E. O’Rourke, Eleanor J. Blitzer et Claire Kremen (09/2011). “A meta-analysis of crop pest and natural enemy response to landscape complexity”. en. In : *Ecology Letters* 14.9, p. 922-932. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2011.01642.x/abstract> (cf. p. 7).
- Clark, James S. (2005). “Why environmental scientists are becoming Bayesians”. In : *Ecology letters* 8.1, p. 2-14. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2004.00702.x/full> (cf. p. 101).
- Codling, Edward A., Michael J. Plank et Simon Benhamou (08/2008). “Random walk models in biology”. en. In : *Journal of The Royal Society Interface* 5.25, p. 813-834. URL : <http://rsif.royalsocietypublishing.org/content/5/25/813> (cf. pp. 65, 100).
- Cushman, Samuel A., Kevin McGarigal et Maile C. Neel (09/2008). “Parsimony in landscape metrics : Strength, universality, and consistency”. In : *Ecological Indicators* 8.5, p. 691-703. URL : <http://www.sciencedirect.com/science/article/pii/S1470160X07001306> (cf. pp. 26, 28, 64).
- DeAngelis, Donald L. et Volker Grimm (06/2014). “Individual-based models in ecology after four decades”. In : *F1000Prime Reports* 6. URL : <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4047944/> (cf. p. 100).
- Duelli, Peter (04/1997). “Biodiversity evaluation in agricultural landscapes : An approach at two different scales”. In : *Agriculture, Ecosystems & Environment*. Biodiversity in Agriculture - for a Sustainable Future 62.2, p. 81-91. URL : <http://www.sciencedirect.com/science/article/pii/S0167880996011437> (cf. pp. 2, 25).
- Dunning, John B., Brent J. Danielson et H. Ronald Pulliam (1992). “Ecological Processes That Affect Populations in Complex Landscapes”. In : *Oikos* 65.1, p. 169-175. URL : <http://www.jstor.org/stable/3544901> (cf. p. 4).
- Edelman, Andrew J. (02/2012). “Positive Interactions between Desert Granivores : Localized Facilitation of Harvester Ants by Kangaroo Rats”. In : *PLoS ONE* 7.2. URL : <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3279350/> (cf. pp. 29, 30, 65).
- Eyre, M. D., M. L. Luff et C. Leifert (01/2013). “Crop, field boundary, productivity and disturbance influences on ground beetles (Coleoptera, Carabidae) in the agroecosystem”. In : *Agriculture, Ecosystems & Environment* 165, p. 60-67. URL : <http://www.sciencedirect.com/science/article/pii/S0167880912004549> (cf. p. 7).
- Fahrig, Lenore (01/1998). “When does fragmentation of breeding habitat affect population survival?” In : *Ecological Modelling* 105.2-3, p. 273-292. URL : <http://www.sciencedirect.com/science/article/pii/S0304380097001634> (cf. p. 59).
- Fahrig, Lenore, Jacques Baudry, Lluís Brotons, Françoise G. Burel, Thomas O. Crist, Robert J. Fuller, Clelia Sirami, Gavin M. Siriwardena et Jean-Louis Martin (02/2011). “Functional

- landscape heterogeneity and animal biodiversity in agricultural landscapes”. en. In : *Ecology Letters* 14.2, p. 101-112. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2010.01559.x/abstract> (cf. pp. 2, 26, 59, 64).
- Fahrig, Lenore, Judith Girard, Dennis Duro, Jon Pasher, Adam Smith, Steve Javorek, Douglas King, Kathryn Freemark Lindsay, Scott Mitchell et Lutz Tischendorf (02/2015). “Farmlands with smaller crop fields have higher within-field biodiversity”. In : *Agriculture, Ecosystems & Environment* 200, p. 219-234. URL : <http://www.sciencedirect.com/science/article/pii/S0167880914005192> (cf. pp. 7, 9, 26, 43, 59, 63).
- Fahrig, Lenore et William K. Nuttle (2005). “Population Ecology in Spatially Heterogeneous Environments”. en. In : *Ecosystem Function in Heterogeneous Landscapes*. Sous la dir. de Gary M. Lovett, Monica G. Turner, Clive G. Jones et Kathleen C. Weathers. DOI : 10.1007/0-387-24091-8_6. Springer New York, p. 95-118. URL : http://link.springer.com/chapter/10.1007/0-387-24091-8_6 (cf. pp. 25, 26).
- Faivre, Robert, Bertrand Iooss, Stéphanie Mahévas, David Makowski et Hervé Monod (04/2013). *Analyse de sensibilité et exploration de modèles : application aux sciences de la nature et de l’environnement*. fr. Editions Quae (cf. pp. 99, 100).
- Forman, Richard T. T. (11/1995a). *Land Mosaics : The Ecology of Landscapes and Regions*. en. Google-Books-ID : sSRNU_5P5nwC. Cambridge University Press (cf. p. 2).
- (06/1995b). “Some general principles of landscape and regional ecology”. en. In : *Landscape Ecology* 10.3, p. 133-142. URL : <https://link.springer.com/article/10.1007/BF00133027> (cf. p. 2).
- Forman, Richard T. T. et Michel Godron (02/1986). *Landscape Ecology*. en. Google-Books-ID : ZvNEVs2MWqcC. Wiley (cf. p. 2).
- Fortin, Marie-Josée et Mark RT Dale (2005). *Spatial analysis : a guide for ecologists*. Cambridge University Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=bTT7tEiTFLoC&oi=fnd&pg=PA1&dq=spatial+analysis:+a+guide+for+ecologists.&ots=qSgTFacVRj&sig=gAXaISwd8Y86GLbd0tn3z67T7vA> (cf. pp. 4, 60).
- Friedman, Jerome H. (04/2001). “Greedy function approximation : A gradient boosting machine.” In : *The Annals of Statistics* 29.5, p. 1189-1232. URL : <http://projecteuclid.org/euclid.aos/1013203451> (cf. p. 68).
- Gardiner, M. M., D. A. Landis, C. Gratton, C. D. DiFonzo, M. O’Neal, J. M. Chacon, M. T. Wayo, N. P. Schmidt, E. E. Mueller et G. E. Heimpel (01/2009). “Landscape diversity enhances biological control of an introduced crop pest in the north-central USA”. In : *Ecological Applications* 19.1, p. 143-154. URL : <http://www.esajournals.org/doi/abs/10.1890/07-1265.1> (cf. p. 7).
- Gaucherel, C. (11/2008). “Neutral models for polygonal landscapes with linear networks”. In : *Ecological Modelling* 219.1–2, p. 39-48. URL : <http://www.sciencedirect.com/science/article/pii/S0304380008003803> (cf. pp. 28, 64).
- Geiger, Flavia *et al.* (03/2010). “Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland”. In : *Basic and Applied Ecology* 11.2, p. 97-105. URL : <http://www.sciencedirect.com/science/article/pii/S1439179109001388> (cf. p. 24).

- Gillespie, Mark A. K. *et al.* (04/2017). “A method for the objective selection of landscape-scale study regions and sites at the national level”. en. In : *Methods in Ecology and Evolution*, n/a-n/a. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12779/abstract> (cf. p. 63).
- Grimm, Volker, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M. Mooij, Steven F. Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand et Donald L. DeAngelis (11/2005). “Pattern-Oriented Modeling of Agent-Based Complex Systems : Lessons from Ecology”. en. In : *Science* 310.5750, p. 987-991. URL : <http://www.sciencemag.org/content/310/5750/987> (cf. pp. 8-10, 56, 60, 96, 100).
- Hall, Linnea S., Paul R. Krausman et Michael L. Morrison (1997). “The Habitat Concept and a Plea for Standard Terminology”. In : *Wildlife Society Bulletin (1973-2006)* 25.1, p. 173-182. URL : <http://www.jstor.org/stable/3783301> (cf. p. 65).
- Hallmann, Caspar A. *et al.* (2017). “More than 75 percent decline over 27 years in total flying insect biomass in protected areas”. eng. In : *PloS One* 12.10, e0185809 (cf. p. 20).
- Hastie, Trevor J. et Robert J. Tibshirani (1990). *Generalized additive models*. T. 43. CRC Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=qa29r1Ze1coC&oi=fnd&pg=PR13&dq=Generalized+Additive+Models+hastie&ots=j32WgmzYpI&sig=10sLwH2xhu0avZiqYA-ELdUo4tg> (cf. p. 69).
- Hastie, Trevor, Robert Tibshirani et Jerome Friedman (08/2009). *The Elements of Statistical Learning : Data Mining, Inference, and Prediction, Second Edition*. en. Google-Books-ID : tVIjmNS3Ob8C. Springer Science & Business Media (cf. p. 68).
- Holland J. M., Begbie M. (2004). “The spatial dynamics and movement of *Pterostichus melanarius* and *P. madidus* (Carabidae) between and within arable fields in the UK”. In : *International Journal of Ecology and Environmental Sciences* 30, p. 35-50 (cf. pp. 24, 61).
- Holland, J. M. et M. L. Luff (06/2000). “The Effects of Agricultural Practices on Carabidae in Temperate Agroecosystems”. en. In : *Integrated Pest Management Reviews* 5.2, p. 109-129. URL : <http://link.springer.com/article/10.1023/A%3A1009619309424> (cf. pp. 59, 61).
- Holland, J. M., C. F. G. Thomas, T. Birkett, S. Southway et H. Oaten (12/2005). “Farm-scale spatiotemporal dynamics of predatory beetles in arable crops”. en. In : *Journal of Applied Ecology* 42.6, p. 1140-1152. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2664.2005.01083.x/abstract> (cf. pp. 21, 64, 67).
- Holling, C. S. et Craig R. Allen (06/2002). “Adaptive Inference for Distinguishing Credible from Incredible Patterns in Nature”. en. In : *Ecosystems* 5.4, p. 319-328. URL : <https://link.springer.com/article/10.1007/s10021-001-0076-2> (cf. pp. 9, 35, 96).
- Holmes, E. E., M. A. Lewis, J. E. Banks et R. R. Veit (01/1994). “Partial Differential Equations in Ecology : Spatial Interactions and Population Dynamics”. en. In : *Ecology* 75.1, p. 17-29. URL : <http://onlinelibrary.wiley.com/doi/10.2307/1939378/abstract> (cf. p. 102).
- Honek, A. (1988). “The effect of crop density and microclimate on pitfall trap catches of Carabidae, Staphylinidae (Coleoptera), and Lycosidae (Araneae) in cereal fields”. In : *Pedobiologia (Jena)* 32.3-4, p. 233-242. URL : <http://www.refdoc.fr/Detailnotice?idarticle> (cf. pp. 24, 101).

- Honek, Alois et Vojtech Jarosik (12/2000). “The role of crop density, seed and aphid presence in diversification of field communities of Carabidae (Coleoptera)”. In : *European Journal of Entomology* 97.4, p. 517-525. URL : <http://www.eje.cz/doi/10.14411/eje.2000.080.html> (cf. p. 59).
- Huzimura, Ryoitiro et Toyoki Matsuyama (12/1999). “A Mathematical Model with a Modified Logistic Approach for Singly Peaked Population Processes”. In : *Theoretical Population Biology* 56.3, p. 301-306. URL : <http://www.sciencedirect.com/science/article/pii/S0040580999914260> (cf. p. 65).
- Illian, Dr Janine, Prof Antti Penttinen, Dr Helga Stoyan et Dr Dietrich Stoyan (04/2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. en. John Wiley & Sons (cf. pp. 28-30).
- J. Haynes, Kyle et James T. Cronin (04/2006). “Interpatch movement and edge effects : the role of behavioral responses to the landscape matrix”. en. In : *Oikos* 113.1, p. 43-54. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.0030-1299.2006.13977.x/abstract> (cf. p. 83).
- Jackson, Heather Bird et Lenore Fahrig (01/2015). “Are ecologists conducting research at the optimal scale?” en. In : *Global Ecology and Biogeography* 24.1, p. 52-63. URL : <http://onlinelibrary.wiley.com/doi/10.1111/geb.12233/abstract> (cf. p. 59).
- Johnson, Jerald B. et Kristian S. Omland (02/2004). “Model selection in ecology and evolution”. In : *Trends in Ecology & Evolution* 19.2, p. 101-108. URL : <http://www.sciencedirect.com/science/article/pii/S0169534703003458> (cf. pp. 35, 60, 69, 83).
- Jongejans, Eelke, Olav Skarpaas et Katriona Shea (03/2008). “Dispersal, demography and spatial population models for conservation and control management”. In : *Perspectives in Plant Ecology, Evolution and Systematics*. Space matters - Novel developments in plant ecology through spatial modelling 9.3-4, p. 153-170. URL : <http://www.sciencedirect.com/science/article/pii/S1433831907000467> (cf. pp. 5, 8, 59, 61, 81).
- Jonsson, Mattias, Riccardo Bommarco, Barbara Ekbom, Henrik G. Smith, Jan Bengtsson, Berta Caballero-Lopez, Camilla Winqvist et Ola Olsson (03/2014). “Ecological production functions for biological control services in agricultural landscapes”. en. In : *Methods in Ecology and Evolution* 5.3, p. 243-252. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12149/abstract> (cf. pp. 5, 59).
- Kaufmann, Leonard et Peter Rousseeuw (01/1987). “Clustering by Means of Medoids”. In : *Data Analysis based on the L1-Norm and Related Methods*, p. 405-416 (cf. p. 38).
- Kiêu, Kiên, Katarzyna Adamczyk-Chauvat, Hervé Monod et Radu S. Stoica (02/2013). “A completely random T-tessellation model and Gibbsian extensions”. In : *arXiv :1302.1809 [math, stat]*. arXiv : 1302.1809. URL : <http://arxiv.org/abs/1302.1809> (cf. p. 40).
- Kromp, Bernhard (06/1999). “Carabid beetles in sustainable agriculture : a review on pest control efficacy, cultivation impacts and enhancement”. In : *Agriculture, Ecosystems & Environment* 74.1-3, p. 187-228. URL : <http://www.sciencedirect.com/science/article/pii/S0167880999000377> (cf. pp. 24, 100).
- Landis, Douglas A., Stephen D. Wratten et Geoff M. Gurr (2000). “Habitat Management to Conserve Natural Enemies of Arthropod Pests in Agriculture”. In : *Annual Review of En-*

- tomology* 45.1, p. 175-201. URL : <http://dx.doi.org/10.1146/annurev.ento.45.1.175> (cf. pp. 5, 24, 59).
- Lausch, Angela, Thomas Blaschke, Dagmar Haase, Felix Herzog, Ralf-Uwe Syrbe, Lutz Tischendorf et Ulrich Walz (01/2015). “Understanding and quantifying landscape structure – A review on relevant process characteristics, data models and landscape metrics”. In : *Ecological Modelling*. Use of ecological indicators in models 295, p. 31-41. URL : <http://www.sciencedirect.com/science/article/pii/S0304380014003974> (cf. p. 28).
- Le Ber, Florence, Claire Lavigne, Katarzyna Adamczyk, Frédérique Angevin, Nathalie Colbach, Jean-François Mari et Hervé Monod (2009). “Neutral modelling of agricultural landscapes by tessellation methods-Application for gene flow simulation.” In : *Ecological Modelling* 220, p. 3536-3545. URL : <https://hal.archives-ouvertes.fr/hal-00409081> (cf. pp. 28, 40, 64).
- Marrec, Ronan, Isabelle Badenhauer, Vincent Bretagnolle, Luca Börger, Marilyn Roncoroni, Nadine Guillon et Bertrand Gauffre (01/2015). “Crop succession and habitat preferences drive the distribution and abundance of carabid beetles in an agricultural landscape”. In : *Agriculture, Ecosystems & Environment* 199, p. 282-289. URL : <http://www.sciencedirect.com/science/article/pii/S016788091400468X> (cf. pp. 7, 9, 24, 33, 67).
- Martin, Emily A., Björn Reineking, Bumsuk Seo et Ingolf Steffan-Dewenter (04/2013). “Natural enemy interactions constrain pest control in complex agricultural landscapes”. In : *Proceedings of the National Academy of Sciences of the United States of America* 110.14, p. 5534-5539. URL : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3619341/> (cf. p. 103).
- Mateu, Jorge et Werner G. Müller (2012). *Spatio-temporal design : advances in efficient data acquisition*. John Wiley & Sons. URL : <https://books.google.fr/books?hl=fr&lr=&id=GWRqLPcTEnc&oi=fnd&pg=PT9&dq=Spatio-temporal+Design:+Advances+in+Efficient+Data+Acquisition&ots=DkjUEB6XE&sig=mLzvK-A0iKU5JtwAJp3tMASAGUM> (cf. p. 60).
- McGarigal, Kevin (2006). “Landscape Pattern Metrics”. en. In : *Encyclopedia of Environmetrics*. John Wiley & Sons, Ltd. URL : <http://onlinelibrary.wiley.com/doi/10.1002/9780470057339.val006.pub2/abstract> (cf. p. 28).
- McIntire, Eliot J. B. et Alex Fajardo (01/2009). “Beyond description : the active and effective way to infer processes from spatial patterns”. In : *Ecology* 90.1, p. 46-56. URL : <http://www.esajournals.org/doi/abs/10.1890/07-2096.1> (cf. pp. 8, 9, 54, 56, 60).
- McIntire, Eliot J. B., Ghislain Rompré et Paul M. Severns (05/2013). “Biased correlated random walk and foray loop : which movement hypothesis drives a butterfly metapopulation?” en. In : *Oecologia* 172.1, p. 293-305. URL : <http://link.springer.com/article/10.1007/s00442-012-2475-9> (cf. pp. 10, 54, 60, 100).
- Moloney, Kirk A. et Simon A. Levin (1996). “The effects of disturbance architecture on landscape-level population dynamics”. In : *Ecology*, p. 375-394. URL : <http://www.jstor.org/stable/2265616> (cf. pp. 5, 59).
- Morris, Max D. (05/1991). “Factorial Sampling Plans for Preliminary Computational Experiments”. In : *Technometrics* 33.2, p. 161-174. URL : <http://www.tandfonline.com/doi/abs/10.1080/00401706.1991.10484804> (cf. pp. 70, 78, 125).

- Nathan, Ran (09/2008). “An emerging movement ecology paradigm”. en. In : *Proceedings of the National Academy of Sciences* 105.49, p. 19050-19051. URL : <http://www.pnas.org/content/105/49/19050> (cf. pp. 10, 54, 60).
- Nisbet, R. M. et W. S. C. Gurney (02/1976). “Population dynamics in a periodically varying environment”. In : *Journal of Theoretical Biology* 56.2, p. 459-475. URL : <http://www.sciencedirect.com/science/article/pii/S0022519376800860> (cf. p. 120).
- Othmani, Imed (05/1998). “Optimisation multicritère : fondements et concepts”. fr. Thèse de doct. Université Joseph-Fourier - Grenoble I. URL : <https://tel.archives-ouvertes.fr/tel-00004900/document> (cf. p. 104).
- Padrón, Víctor et María Cristina Trevisan (02/2006). “Environmentally induced dispersal under heterogeneous logistic growth”. In : *Mathematical Biosciences* 199.2, p. 160-174. URL : <http://www.sciencedirect.com/science/article/pii/S0025556405002075> (cf. pp. 34, 65).
- Papaïx, Julien, Suzanne Touzeau, Hervé Monod et Christian Lannou (07/2014). “Can epidemic control be achieved by altering landscape connectivity in agricultural systems?” In : *Ecological Modelling* 284, p. 35-47. URL : <http://www.sciencedirect.com/science/article/pii/S0304380014001951> (cf. p. 99).
- Parisey, Nicolas, Yoann Bourhis, Lionel Roques, Samuel Soubeyrand, Benoît Ricci et Sylvain Poggi (12/2016). “Rearranging agricultural landscapes towards habitat quality optimisation : In silico application to pest regulation”. In : *Ecological Complexity* 28, p. 113-122. URL : <https://www.sciencedirect.com/science/article/pii/S1476945X16300538> (cf. pp. 5, 34, 59, 101).
- Peck, Steven L. (10/2004). “Simulation as experiment : a philosophical reassessment for biological modeling”. eng. In : *Trends in Ecology & Evolution* 19.10, p. 530-534 (cf. pp. 8-10, 96, 100).
- Platt, J. R. (10/1964). “Strong Inference : Certain systematic methods of scientific thinking may produce much more rapid progress than others”. eng. In : *Science (New York, N.Y.)* 146.3642, p. 347-353 (cf. pp. 56, 60, 83).
- Press, William H. (2007). *Numerical recipes 3rd edition : The art of scientific computing*. Cambridge university press. URL : <https://books.google.fr/books?hl=fr&lr=&id=1aA0dzK3FegC&oi=fnd&pg=PA1&dq=press+numerical+recipes&ots=3jRmG8Grog&sig=C3unnUCUA2krwf0mzPVkAs1LcKY> (cf. p. 37).
- Puech, Camille (12/2014). “Hétérogénéité des pratiques agricoles biologiques et conventionnelles dans les paysages bretons : effets sur les communautés d’insectes auxiliaires à différentes échelles”. Thèse de doct. Rennes 1. URL : <http://www.theses.fr/2014REN1S117> (cf. p. 97).
- Puech, Camille, Sylvain Poggi, Jacques Baudry, Jacques et Stéphanie Aviron Stéphanie (10/2014). “Do farming practices affect natural enemies at the landscape scale?” In : *Landscape Ecology* 30 (cf. p. 7).
- Retho, Benjamin, Cédric Gaucherel et Pablo Inchausti (03/2008). “Modeling spatially explicit population dynamics of *Pterostichus melanarius* H11. (Coleoptera : Carabidae) in response to changes in the composition and configuration of agricultural landscapes”. In : *Landscape and Urban Planning* 84, p. 191-199. URL : <http://www.sciencedirect.com/passerelle.univ-rennes1.fr/science/article/pii/S0169204607001818> (cf. p. 120).

- Ricci, Benoît, Pierre Franck, Muriel Valantin-Morison, David A. Bohan et Claire Lavigne (09/2013). “Do species population parameters and landscape characteristics affect the relationship between local population abundance and surrounding habitat amount?” In : *Ecological Complexity* 15, p. 62-70. URL : <http://www.sciencedirect.com/science/article/pii/S1476945X13000330> (cf. p. 8).
- Rijnsdorp, A. D. (05/1980). “Pattern of movement in and dispersal from a dutch forest of *Carabus problematicus* Hbst. (Coleoptera, Carabidae)”. eng. In : *Oecologia* 45.2, p. 274-281 (cf. p. 24).
- Roques, Lionel, Marie-Anne Auger-Rozenberg et Alain Roques (11/2008). “Modelling the impact of an invasive insect via reaction-diffusion”. In : *Mathematical Biosciences* 216.1, p. 47-55. URL : <http://www.sciencedirect.com/science/article/pii/S0025556408001235> (cf. pp. 33, 65).
- Roques, Lionel et M. D. Chekroun (03/2010). “Does reaction-diffusion support the duality of fragmentation effect?” In : *Ecological Complexity* 7.1, p. 100-106. URL : <http://www.sciencedirect.com/science/article/pii/S1476945X09000646> (cf. p. 65).
- Roques, Lionel et Samuel Soubeyrand (2013). *Modèles de réaction-diffusion pour l'écologie spatiale*. français. Versailles, France : Ed. Quae, impr. 2013 (cf. pp. 33, 34).
- Rousseeuw, Peter J. (11/1987). “Silhouettes : A graphical aid to the interpretation and validation of cluster analysis”. In : *Journal of Computational and Applied Mathematics* 20.Supplement C, p. 53-65. URL : <http://www.sciencedirect.com/science/article/pii/0377042787901257> (cf. p. 38).
- Savage, David et Michael Renton (01/2014). “Requirements, design and implementation of a general model of biological invasion”. In : *Ecological Modelling* 272, p. 394-409. URL : <http://www.sciencedirect.com/science/article/pii/S0304380013004638> (cf. pp. 5, 59).
- Schröder, B. et R. Seppelt (12/2006). “Analysis of pattern–process interactions based on landscape models—Overview, general concepts, and methodological issues”. In : *Ecological Modelling. Pattern and Processes of Dynamic Mosaic Landscapes – Modelling, Simulation, and Implications* 199.4, p. 505-516. URL : <http://www.sciencedirect.com/science/article/pii/S0304380006002936> (cf. pp. 3, 4, 59-61, 96).
- Schweiger, Andreas H., Severin D. H. Irl, Manuel J. Steinbauer, Jürgen Dengler et Carl Beierkuhnlein (04/2016). “Optimizing sampling approaches along ecological gradients”. en. In : *Methods in Ecology and Evolution* 7.4, p. 463-471. URL : <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12495/abstract> (cf. pp. 10, 60, 101).
- Shigesada, Nanako et Kohkichi Kawasaki (02/1997). *Biological Invasions : Theory and Practice*. en. Oxford University Press, UK (cf. pp. 33, 65).
- Soubeyrand, Samuel et Lionel Roques (04/2014). “Parameter estimation for reaction-diffusion models of biological invasions”. en. In : *Population Ecology* 56.2, p. 427-434. URL : <http://link.springer.com/10.1007/s10144-013-0415-0> (cf. p. 101).
- Sunderland, K.D. (01/2002). “Invertebrate pest control by carabids”. In : *The Agroecology of Carabid Beetles*, p. 165-214 (cf. p. 24).
- Teillard, Félix (10/2012). “Reconciling food production and biodiversity in farmlands : the role of agricultural intensity and its spatial allocation”. In : (cf. p. 104).

- Thiele, Hans Ulrich (1977). *Carabid beetles in their environments : a study on habitat selection by adaptations in physiology and behaviour*. en. Google-Books-ID : bK_wAAAAMAAJ. Springer-Verlag (cf. p. 24).
- Thomas, C. F. G., N. J. Brown et D. A. Kendall (04/2006). “Carabid movement and vegetation density : Implications for interpreting pitfall trap data from split-field trials”. In : *Agriculture, Ecosystems & Environment* 113.1–4, p. 51-61. URL : <http://www.sciencedirect.com/science/article/pii/S0167880905004330> (cf. p. 24).
- Tilman, David et Peter M. Kareiva (1997). *Spatial ecology : the role of space in population dynamics and interspecific interactions*. T. 30. Princeton University Press. URL : <https://books.google.fr/books?hl=fr&lr=&id=QdqC8PvWB7AC&oi=fnd&pg=PR7&dq=Spatial+ecology:+the+role+of+space+in+population+dynamics+and+interspecific+interactions.&ots=bKYtjihHzj&sig=4VpUdp00mt23W37oQZ6aRi7BJAY> (cf. pp. 8, 60).
- Tuda, Midori (03/2007). “Understanding mechanism of spatial ecological phenomena : a preface to the special feature on “Spatial statistics””. en. In : *Ecological Research* 22.2, p. 183-184. URL : <http://link.springer.com/article/10.1007/s11284-006-0328-x> (cf. pp. 8, 60).
- Turner, Monica G. (2005). “LANDSCAPE ECOLOGY : What Is the State of the Science?” In : *Annual Review of Ecology, Evolution, and Systematics* 36.1, p. 319-344. URL : <http://dx.doi.org/10.1146/annurev.ecolsys.36.102003.152614> (cf. p. 59).
- Turner, Monica G. et Robert H. Gardner (2015). *Landscape Ecology in Theory and Practice*. en. New York, NY : Springer New York. URL : <http://link.springer.com/10.1007/978-1-4939-2794-4> (cf. pp. 2, 6-8).
- Vasseur, Chloé, Alexandre Joannon, Stéphanie Aviron, Françoise Burel, Jean-Marc Meynard et Jacques Baudry (02/2013). “The cropping systems mosaic : How does the hidden heterogeneity of agricultural landscapes drive arthropod populations?” In : *Agriculture, Ecosystems & Environment*. Landscape ecology and biodiversity in agricultural landscapes 166, p. 3-14. URL : <http://www.sciencedirect.com/science/article/pii/S0167880912003246> (cf. pp. 2, 25).
- Vialatte, Aude, Noëlline Tsafack, Diab Al Hassan, Rémi Dufflot, Manuel Plantegenest, Annie Ouin, Johanna Villenave-Chasset et Aude Ernoult (12/2016). “Landscape potential for pollen provisioning for beneficial insects favours biological control in crop fields”. English. In : *Landscape Ecology* 3.32, p. 465-480. URL : https://www.infona.pl/resource/bwmeta1.element.springer-doi-10_1007-S10980-016-0481-8 (cf. p. 103).
- Vinatier, Fabrice (01/2010). *Dynamique spatiale du charançon du bananier en interaction avec le système de culture et l'organisation paysagère*. AgroParisTech. URL : <http://www.theses.fr/2010AGPT0068> (cf. p. 4).
- Vinatier, Fabrice, Philippe Tixier, Pierre-François Duyck et Françoise Lescourret (2011). “Factors and mechanisms explaining spatial heterogeneity : a review of methods for insect populations”. en. In : *Methods in Ecology and Evolution* 2.1, p. 11-22. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2010.00059.x/abstract> (cf. pp. 9, 81, 100).
- Wallin, H. et B. S. Ekbom (10/1988). “Movements of carabid beetles (Coleoptera : Carabidae) inhabiting cereal fields : a field tracing study”. en. In : *Oecologia* 77.1, p. 39-43. URL : <http://link.springer.com/article/10.1007/BF00380922> (cf. p. 24).

- Wallin, Henrik et Barbara Ekblom (10/1994). "Influence of Hunger Level and Prey Densities on Movement Patterns in Three Species of Pterostichus Beetles (Coleoptera : Carabidae)". en. In : *Environmental Entomology* 23.5, p. 1171-1181. URL : <http://ee.oxfordjournals.org/content/23/5/1171> (cf. p. 24).
- Wiens, John A. (1999). "The Science and Practice of Landscape Ecology". en. In : *Landscape Ecological Analysis*. DOI : 10.1007/978-1-4612-0529-6_16. Springer, New York, NY, p. 371-383. URL : https://link.springer.com/chapter/10.1007/978-1-4612-0529-6_16 (cf. p. 3).
- Winder, L., G. J. K. Griffiths, J. N. Perry, C. J. Alexander, J. M. Holland, P. J. Kennedy et A. Birt (12/2005). "The role of large-scale spatially explicit and small-scale localized processes on the population dynamics of cereal aphids". eng. In : *Bulletin of Entomological Research* 95.6, p. 579-587 (cf. p. 24).
- Winsberg, Eric (01/2003). "Simulated Experiments : Methodology for a Virtual World". In : *Philosophy of Science* 70.1, p. 105-125. URL : <http://www.journals.uchicago.edu/doi/abs/10.1086/367872> (cf. p. 10).
- Wood, Simon N. (02/2003). "Thin plate regression splines". en. In : *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 65.1, p. 95-114. URL : <http://onlinelibrary.wiley.com/doi/10.1111/1467-9868.00374/abstract> (cf. p. 69).
- Yang, Liu (2006). *Distance Metric Learning : A Comprehensive Survey* (cf. p. 102).
- Zurell, Damaris *et al.* (04/2010). "The virtual ecologist approach : simulating data and observers". en. In : *Oikos* 119.4, p. 622-635. URL : <http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0706.2009.18284.x/abstract> (cf. pp. 10, 54, 60).

Appendices



Annexes relatives au chapitre 2

B

Annexes relatives à l'article

B.1 Process models calibration

Per capita growth rate was calculated from oviposition rates found in the literature (Retho *et al.* 2008). Population densities $P(\mathbf{x}, t)$ are expressed as a ratio of the maximum carrying capacity so that results can be generalised to other species (Nisbet *et Gurney* 1976). Carrying capacities for the *PM*s with inhomogeneous reaction (K_1, K_2 and K_3) were set to 0.1, 0.55 and 1 (respectively "low", "mid" and "high" quality habitat). The other reaction parameters were set such as the integral of carrying capacity along time at the landscape scale equates the integral for the inhomogeneous reaction.

Thus we have for K , the carrying capacity for the model $\mathcal{M}_{HV,IC}$

$$K = p_1 K_1 + p_2 K_2 + p_3 K_3$$

which is

$$K = \frac{K_1 + K_2 + K_3}{3}$$

as we constrained equal proportion of the three landcover types for all landscapes. Similarly we have for β_r , the exponential decrease parameter for the model $\mathcal{M}_{HV,IC}$

$$\int_0^T K_0 e^{-\beta_r t} dt = \frac{K_1 + K_2 + K_3}{3}$$

$$\Leftrightarrow \left[\frac{-e^{-\beta_r t}}{\beta_r} \right]_0^T = \frac{K_1 + K_2 + K_3}{3K_0}$$

$$\Leftrightarrow \frac{1 - e^{-\beta_r T}}{\beta_r} = \frac{K_1 + K_2 + K_3}{3K_0}$$

with $T = 1$, being

$$\Leftrightarrow \frac{1 - e^{-\beta_r}}{\beta_r} = \frac{K_1 + K_2 + K_3}{3K_0}$$

which is numerically approached when a value is set for K_0 ($K_0 = 0.9$).

Diffusion parameters for the *PMs* with inhomogeneous diffusion (μ_1 , μ_2 and μ_3) were taken as the minimum, maximum and mean values of *P. melanarius* found in [Allema, Werf, Groot *et al.* \(2015\)](#) scaled to the extent of our spatial domains. As before, the diffusion parameters for the other *PMs* were derived from the integral in time and space of these values.

B.2 Numerical scheme

Space and time were scaled down to unity, *i.e.* we consider the spatial domain as the unit square with side length of 1 unit space and the simulation duration as 1 unit time. For the numerical calculation, the spatial domain is discretised and characterised by a matrix of land-cover dependant values. The spatial resolution is 0.01×0.01 space unit, defining a 100×100 cells grid. Time is discretised in 1000 time steps. The reaction diffusion model is solved by operator splitting with the demographic component solved by a Runge-Kutta 4 explicit numerical scheme and the diffusive component solved with a finite volume method and an Alternate Direction Implicit numerical scheme, as described in [Bourhis, Poggi, Mammeri, Cortesero *et al.* \(2015\)](#). The matrix of initial densities $P(\mathbf{x}, 0)$ equals half the maximum carrying capacity in each cell ($P(\mathbf{x}, 0) = 0.5\forall\mathbf{x}$).

Each simulation of the model \mathcal{M}_j on the domain Ω_j generates a matrix $A_{\mathcal{M}_i, \Omega_j}$ of dimension $100 \times 100 \times 1000$.

B.3 Population metrics computed after spatio-temporal sampling

FIGURE B.1 – Table 5 : Population metrics computed after sampling

Name	Description	Formula
$\phi_{mean.t.mean.sp}$	mean in time of mean population in space	$= \frac{\sum_{t=1}^T \bar{P}_t}{T}$
$\phi_{mean.t.sd.sp}$	mean in time of population standard deviation in space	$= \frac{\sum_{t=1}^T \sqrt{\frac{1}{s} \sum_{i=1}^s (P_{i,t} - \bar{P}_t)^2}}{T}$
$\phi_{mean.t.min.sp}$	mean in time of population minima in space	$= \frac{\sum_{t=1}^T m_t}{T}$ with $m_t = \min P_i(t)$
$\phi_{mean.t.max.sp}$	mean in time of population maxima in space	$= \frac{\sum_{t=1}^T m_t}{T}$ with $m_t = \max P_i(t)$
$\phi_{sd.t.mean.sp}$	standard deviation in time of mean population in space	$= \sqrt{\frac{1}{T} \sum_{t=1}^T (\bar{P}_t - \bar{P})^2}$
$\phi_{sd.t.sd.sp}$	standard deviation in time of population standard deviation in space	$= \frac{\sum_{j=1}^T \sqrt{\frac{1}{s} \sum_{t=1}^T (\sigma_t - \bar{\sigma})^2}}{T}$ with $\sigma_t = \sqrt{\frac{1}{s} \sum_{i=1}^s (P_{i,t} - \bar{P}_t)^2}$ and $\bar{\sigma} = \frac{\sum_{t=1}^T \sigma_t}{T}$
$\phi_{sd.t.min.sp}$	standard deviation in time of population minima in space	$= \sqrt{\frac{1}{T} \sum_{t=1}^T (m_t - \bar{m})^2}$ with $m_t = \min P_i(t)$
$\phi_{sd.t.max.sp}$	standard deviation in time of population maxima in space	$= \sqrt{\frac{1}{T} \sum_{t=1}^T (m_t - \bar{m})^2}$ with $m_t = \max P_i(t)$

FIGURE B.2 – Table 6 : Population metrics computed after sampling (continued)

Name	Description	Formula
$\phi_{mean.sp.sd.t}$	mean in space of population standard deviation in time	$= \frac{1}{S} \sum_{s=1}^S \sqrt{\frac{1}{T} \sum_{i=1}^T (P_{i,s} - \bar{P}_s)^2}$
$\phi_{mean.sp.min.t}$	mean in space of population minima in time	$= \frac{\sum_{s=1}^S m_s}{T}$ with $m_t = \min P_i(s)$
$\phi_{mean.sp.max.t}$	mean in space of population maxima in time	$= \frac{\sum_{s=1}^S m_s}{S}$ with $m_t = \max P_i(S)$
$\phi_{sd.sp.mean.t}$	standard deviation in space of mean population in time	$= \sqrt{\frac{1}{S} \sum_{s=1}^S (\bar{P}_s - \bar{P})^2}$
$\phi_{sd.sp.sd.t}$	standard deviation in space of population standard deviation in time	$= \frac{1}{S} \sum_{j=1}^S \sqrt{\frac{1}{T} \sum_{s=1}^S (\sigma_s - \bar{\sigma})^2}$ with $\sigma_s = \sqrt{\frac{1}{T} \sum_{i=1}^T (P_{i,s} - \bar{P}_s)^2}$ and $\bar{\sigma} = \frac{\sum_{s=1}^S \sigma_s}{S}$
$\phi_{sd.sp.min.t}$	standard deviation in space of population minima in time	$= \sqrt{\frac{1}{S} \sum_{s=1}^S (m_s - \bar{m})^2}$ with $m_s = \min P_i(s)$
$\phi_{sd.sp.max.t}$	standard deviation in space of population maxima in time	$= \sqrt{\frac{1}{S} \sum_{s=1}^S (m_s - \bar{m})^2}$ with $m_s = \max P_i(s)$

B.4 Global sensitivity analysis

Let $y^* = PM_m(v_1^*, \dots, v_p^*)$ and $y' = PM_m(v'_1, \dots, v'_p)$ be the vectors combining the sampled observations of the population dynamics generated by these models, and

Φ^* and Φ' their respective descriptors. We want to assess the effect of the divergence between v^* and v' on the loss of classification performances.

We defined a matrix $V' = (v'_{ij})$ with $1 \leq i \leq n$ and $1 \leq j \leq p$ with p the number of parameters and n the number of rows. We used the Morris method (Morris 1991) to design the parameter space matrix V' , using a one-at-a-time factorial design with 10 levels for each parameter, a grid jump δ of $10/2 = 5$, and $H = 30$ trajectories, leading to a table of $n = 420$ parameter sets.

We simulated the process models and extracted the associated population data, each was sampled with $L = 10$ instances of the optimal sampling design \mathcal{S}^* to give the observation matrix $Y' = (y'_{j,m,l})$

$$y'_{jlm} = PM_m(v'_j) \sim \mathcal{S}_i^*; j = 1, \dots, n; l = 1, \dots, 10; m = 1, \dots, M \quad (\text{B.1})$$

and extracted the associated descriptors Φ'_{jkl} .

Using the reference classifier \mathcal{C}^* , we computed the rate of correct classification for each row of the matrix V' , as

$$A'_j = \frac{1}{M} \frac{1}{L} \sum_{m=1}^M \sum_{l=1}^L (\widehat{y'_{jlm}} = PM_m) \quad (\text{B.2})$$

The loss of classification performance for the PM_m is then defined for each row of V' as

$$G_j = A^* - A'_j \quad (\text{B.3})$$

with $A^* = \frac{1}{L} \frac{1}{M} \sum_{l=1}^L \sum_{m=1}^M (\widehat{y_m^*} = PM_m)$ the rate of correct classifications obtained with the reference parameter set V^* .

We then computed the absolute mean (μ^*) and variance (σ) of the parameters elementary effects as :

$$\mu_i^* = \frac{1}{H} \sum_{h=1}^H |\Delta_i^{(h)} G_j| \quad (\text{B.4})$$

$$\sigma_i = \sqrt{\frac{1}{H-1} \sum_{h=1}^H (\Delta_i^{(h)} G_j - \mu_i)^2} \quad (\text{B.5})$$

with $\Delta_i^{(h)} G_j$ the elementary effect of parameter v_i on the loss of accuracy G_j in the trajectory h of the Morris plan, *i.e.* the variation of G given the parameter variation δv_i .

