

Faculté des bioingénieurs

Cartographie à haute résolution des populations d'oiseaux nicheurs en Brabant wallon

Ecotopes, régression de Poisson et perspectives d'étude

Auteur : Kaye Marvin

Promoteur(s) : Bogaert Patrick et Radoux Julien

Lecteur(s) : Paquet Jean-Yves et Ponette Quentin

Année académique 2021-2022

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de
Bioingénieur : Aménagement du territoire

Remerciements

Je suis particulièrement reconnaissant envers mes deux promoteurs, Monsieur Bogaert et Monsieur Radoux, qui m'ont permis de faire ce mémoire une réussite. Leurs remarques et leurs conseils avisés ont été d'une aide inégalable.

Je tiens à remercier également Monsieur Derouaux pour le partage des données d'Aves et tous les observateurs qui ont contribué aux données recensées sur le site Observations.be qui se sont avérées très utiles pour l'étude.

Je remercie mon jury, Monsieur Paquet et Monsieur Ponette, qui ont accepté de lire et évaluer ce travail.

Enfin, je remercie spécialement tous mes proches qui m'ont soutenu et encouragé durant cette période, avec une mention particulière pour ma compagne pour son aide précieuse dans la relecture de ce mémoire.

Table des matières

Introduction	1
1 Cadre du mémoire	1
2 État de l’art	3
2.1 Les indicateurs de biodiversité	3
2.2 Les oiseaux comme indicateurs de biodiversité	4
2.2.1 Oiseaux et territoires diversifiés	5
2.2.2 Espèces spécialistes et généralistes	5
2.3 Les méthodes de suivis pour la modélisation des oiseaux	6
2.4 Les programmes de suivis d’oiseaux	8
2.5 Aves, pôle ornithologique de Natagora	9
2.5.1 Les relevés de données oiseaux	10
2.6 Les projets LifeWatch	11
2.6.1 Les écotopes	12
2.7 Prédire l’abondance des espèces	15
2.7.1 Aperçu des modèles Espèces-Habitats	15
2.7.2 L’importance de travailler à haute résolution	17
2.7.3 Comparaison entre différents modèles statistiques de régression	18
2.7.4 La régression de Poisson et le modèle linéaire généralisé	19
3 Objectifs du travail	20
4 Plan du travail	20
Matériels et méthode	22
5 Sites d’étude : les carrés kilométriques	22
6 Choix des espèces observées	23
6.1 Oiseaux du milieu urbain	23
6.2 Oiseaux du milieu ouvert et agricole	24
6.3 Oiseaux du milieu forestier	25
6.4 Oiseaux du milieu aquatique et des plans d’eau	26

7	Recueil de bases de données	27
7.1	Observations.be	27
7.2	Relevés de terrain	28
8	Traitement des données	29
8.1	Méthodes d'analyse pour la prédiction des densités des populations	33
9	Validation des résultats	33
10	Analyse de sensibilité	35
10.1	Utilisation des données issues de la campagne de terrain uniquement sans les données provenant du site Observations.be	35
10.2	Ajout des données d'absence	35
10.3	Utilisation d'une variable de référence différente lors du problème de colinéarité entre variables	36
10.4	Utilisation d'un autre modèle de type "Quasi-Poisson"	36
10.5	Caractérisation d'une zone de territoire par groupes d'individus lors du calcul du R^2	36
	Résultats	38
11	Caractéristiques générales des données	38
12	Cartographies représentatives de la réalité de terrain	40
13	Cartographies non représentatives de la réalité de terrain	41
14	Détermination de l'intercept	42
15	Catégorisation des résultats sous forme d'une clé dichotomique	43
16	Tests de sensibilité	44
16.1	Utilisation des données de terrain uniquement	44
16.2	Ajout des données d'absence	44
16.3	Utilisation d'une variable de référence différente	45
16.4	Modélisation avec une régression Quasi-Poisson	46
16.5	Caractérisation d'une zone de territoire par groupe d'individus	47
	Discussion	49

17 Qualité des données issues de l’effort collectif et de terrain	49
18 Répartition spatio-temporelle des données	50
19 Modèle et méthode employée	51
19.1 Représentativité de l’échantillonnage	51
19.2 Précision spatiale des données	52
19.3 Polygones et écotopes	53
19.4 Le modèle	54
19.4.1 Utilisation des données issues de la campagne de terrain uniquement	55
19.4.2 Ajout des données d’absence	56
19.4.3 Dépendance de la variable de référence	56
19.4.4 Utilisation d’une régression de type Quasi-Poisson	57
19.4.5 Méthode d’agrégation	58
20 Perspectives d’étude	59
 Conclusion	 60
Table des annexes	68
Annexe A Schéma explicatif du filtre de Kalman.	68
Annexe B Fiche de terrain - Échantillonnage oiseaux nicheurs 2015-2017.	68
Annexe C Interface de l’application mobile <i>iObs</i>.	69
Annexe D Localisation des chaînes de points d’écoute en Région wallonne. Les couleurs représentent le nombre d’années de suivi (Antoine Derouaux, 2021).	 69
Annexe E Distribution régulière des carrés kilométriques sur tout le territoire wallon (Aves, 2015).	 70
Annexe F Classification des différents écotopes présents dans la zone d’étude selon la majorité des proportions d’occupation du sol.	 71
Annexe G Oiseaux du milieu ouvert et agricole.	72
Annexe H Oiseaux du milieu forestier.	72

Annexe I Oiseaux du milieu urbain.	73
Annexe J Oiseaux du milieu aquatique et des plans d'eau.	73
Annexe K Colonnes et identifiants des données récoltées par Aves - Observation.be.	74
Annexe L Colonnes du fichier Excel des données de terrain téléchargées depuis Observation.be.	75
Annexe M Code de validation des données utilisées par Observations.be.	76
Annexe N Répartition de l'abondance du Pinson des arbres par écotope.	76
Annexe O Répartition non représentative de la réalité de terrain de l'abondance de six autres espèces.	77

Table des figures

1	Living Planet Index (LAMBERTINI, 2020).	1
2	”Carte présentant la densité relative du Rouge-gorge familial (<i>Erithacus rubecula</i>) en Wallonie en 2001-2007, réalisée grâce aux échantillonnages kilométriques. Cette espèce est présente partout, mais montre de fortes variations de densité au travers du territoire, essentiellement liées à la couverture forestière” (AVES, 2015).	11
3	Illustration des modèles conceptuels d’une forêt méditerranéenne dans le sud de la France. Dans ce paysage couvert par une image Ikonos, les arbres sont des entités spatiales détectables individuellement (étoiles sur l’image de gauche). Au centre, la segmentation automatisée des images crée des régions spatiales qui correspondent à l’ancien pare-feu et à la forêt non perturbée composée d’arbres, d’arbustes et de végétation herbacée. À droite, une représentation continue du champ met en évidence la densité de végétation plus faible (gris plus foncé) dans l’ancien pare-feu (JULIEN RADOUX & PATRICK BOGAERT, 2017).	13
4	Affichage des attributs d’un écotope avec les proportions de classes d’occupation du sol.	14
5	Influence de la taille des pixels sur la déviance expliquée, où chaque courbe correspond à une espèce d’oiseau différente. En moyenne, pour toutes les espèces, la déviance expliquée diminue de 41.3 (à 1 mètre de résolution) à 14.4 (à 1000 mètres de résolution) (GOTTSCALK et al., 2011).	17
6	Délimitation de la surface d’étude de 13x13 km (carré noir en surbrillance) et localisation des 14 carrés visités (en bleu) pour l’échantillonnage (AVES, 2015).	22
7	Nombre d’observations encodées annuellement depuis 2008 sur la plateforme Observations.be, consulté en juillet 2022.	28
8	Succession des différentes étapes d’agrégation des abondances à 1 kilomètre de résolution. 1) Calcul de l’abondance par polygone (délimités en noir) proportionnellement à la surface de chaque polygone ; 2) Calcul de la somme des abondances au sein des carrés kilométriques (délimités en rouge).	34
9	Succession des différentes étapes d’agrégation des abondances à 1 kilomètre de résolution. 1) Calcul de l’abondance par polygone (délimités en noir) proportionnellement à la surface de chaque polygone ; 2) Détermination des zones de territoire par la fusion de polygones conjoints ayant la même gamme d’abondance ; 3) Calcul de la somme des abondances au sein des carrés kilométriques (délimités en rouge).	37

10	Distribution des observations enregistrées sur la plateforme <i>Observations.be</i> , ainsi que les relevés de terrain au cours des cinq premiers mois de l'année 2022.	38
11	Répartition de l'abondance de la Mésange charbonnière par écotope.	40
12	Répartition de l'abondance de l'Alouette des champs par écotope.	41
13	Clé dichotomique catégorisant les différents résultats obtenus, où x est le nombre d'individus enregistrés et R^2 est le résultat de la validation.	43
14	Amélioration de la calibration du modèle par l'ajout de données d'absence. A. Cartographie sans les données d'absence ($R^2 = 0,23$); B. Cartographie avec les données d'absence ($R^2 = 0,32$).	45
15	Influence du choix de la variable de référence sur les résultats (cas du Pinson des arbres). A. Plowed250 est choisi comme variable de référence ($R^2 = 0,25$); B. BroadLV250 est choisi comme variable de référence ($R^2 = 0,18$); C. WOpen250 est choisi comme variable de référence ($R^2 = 0,15$); D. COpen250 est choisi comme variable de référence ($R^2 = 0,18$).	46
16	Répartition de l'abondance de la Mésange bleue par écotope par une distribution "Quasi-Poisson".	47
17	Différentes légendes appliquées sur le même paysage. L'image centrale est une représentation basée sur les pixels avec une légende de la règle de la majorité. Deux autres représentations sont construites en cohérence avec cette classification. À gauche, les classes sont basées sur un ensemble de seuils dérivés de règles de type LCCS. À droite, une légende basée sur la majorité des cellules est utilisée (JULIEN RADOUX & PATRICK BOGAERT, 2017).	54

Liste des tableaux

1	Critères d'un bon indicateur selon GREGORY ET AL. (2005).	4
2	Nombre d'espèces en fonction des tendances de leur population wallonne, par habitats (ANTOINE DEROUAUX, 2021).	7
3	Liste des espèces du milieu urbain.	23
4	Liste des espèces du milieu ouvert et agricole.	24
5	Liste des espèces du milieu forestier.	25
6	Liste des espèces du milieu aquatique et des plans d'eau.	26
7	Liste des variables de chaque espèce.	32
8	Nombre d'observations des espèces dont le résultat est représentatif de la réalité de terrain et leur validation cartographique.	40
9	Nombre d'observations des espèces dont le résultat est non représentatif de la réalité de terrain et leur validation cartographique.	41
10	Nombre d'observations des espèces dont le modèle identifie l'intercepte. Le "-" signifie que la validation n'a pas été réalisée car aucune cartographie n'a été obtenue.	42
11	Comparaison entre les méthodes de validation. La validation initiale somme l'abondance estimée par écotope, alors que la validation finale somme l'abondance estimée par zone de territoire définie par carré kilométrique.	48

Liste des acronymes

ABV : Algemene Broedvogels Vlaanderen
AIC : Akaike information criterion
DHEO : Dénombrements Hivernaux d'Oiseaux d'Eau
DDFL : Data-Driven Fuzzy Logic model
GAM : Generalized Additive Model
GEOBIA : Geographic Object-Based Image Analysis
GLM : Generalized Linear Model
HDS : Hierarchical Distance Sampling
HM : Hurdle Model
LPI : Living Planet Index
NBR : Negative Binomial Regression
PECBMS : Pan-European Common Bird Monitoring Scheme
PCM : Preference Curve Model
PRM : Poisson Regression Model
QPR : Quasi-Poisson Regression
RNOB : Réserves naturelles et ornithologiques de Belgique
ROC : Receiver Operating Characteristic
SOCBRU : Surveillance des oiseaux communs de Bruxelles
SOCWAL : Programme du suivi des oiseaux communs en Wallonie
TSS : True Skill Statistics
UCLouvain : Université Catholique de Louvain-la-Neuve
ZIP : Zero-Inflated Poisson

Introduction

1 Cadre du mémoire

Depuis la révolution industrielle, les activités anthropiques se sont considérablement développées partout dans le monde, en détruisant régulièrement les milieux qui les entourent. La perte de ces milieux est en lien direct avec une perte de biodiversité. Pour mieux quantifier cette détérioration, le dernier rapport du WWF fait référence au LPI (pour *Living Planet Index*) qui suit l'abondance de près de 21 000 populations de mammifères, oiseaux, poissons, reptiles et amphibiens dans le monde. Cet indice indique un déclin de 68 % (Figure 1) au sein des populations entre 1970 et 2016 (LAMBERTINI, 2020). Plus précisément, selon le rapport du 7 octobre 2021 publié par BirdLife international à propos de l'avifaune européenne, un déclin de 17 à 19 % est estimé concernant l'abondance globale des oiseaux nicheurs depuis 1980, ce qui se traduit par une perte de 560 à 620 millions d'individus (BURNS et al., 2021).

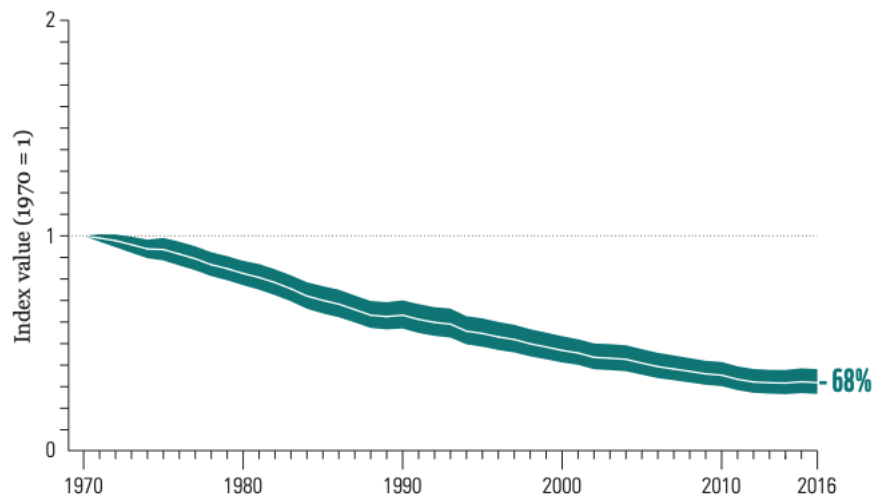


FIGURE 1 – Living Planet Index (LAMBERTINI, 2020).

La Wallonie ne fait pas exception à cette tendance puisqu'un déclin généralisé des populations d'oiseaux sur son territoire y est observé. L'indice d'abondance annuel est calculé sur base de suivis annuels par échantillonnage standardisé pour chaque espèce. Les tendances récentes ne sont pas optimistes : en effet, ces 28 dernières années, une majorité d'espèces communes sont en déclin en Wallonie. Plus précisément, toutes espèces confondues, on observe en moyenne 1% de perte de leurs effectifs par an depuis les années 1990, avec une accélération de cette diminution ces dix dernières années. Les oiseaux inféodés au milieu agricole sont les plus

touchés en raison des modifications des pratiques et de l'intensification de l'agriculture, responsables de la dégradation du milieu (DEROUAUX & PAQUET, 2018).

Au vu de ce déclin, la conservation et la planification de la gestion des espèces est essentielle. Cette gestion exige de connaître les aires de distribution potentielle des espèces et les facteurs environnementaux conditionnant ces répartitions (MOUKRIM et al., 2018). De ce fait, lors de l'identification des éléments structurant le paysage, travailler à haute résolution permet de distinguer au maximum tous ces éléments et donc de mieux comprendre les relations entre les espèces et leur habitat. Dès lors, des modèles mathématiques tentent de quantifier la relation entre les caractéristiques environnementales et l'abondance d'une espèce. Le choix d'utiliser un modèle plutôt qu'un autre est régi par la connaissance de l'espèce et les caractéristiques des données disponibles (POTTS & ELITH, 2006).

En Wallonie, ce déclin a été observé et mesuré grâce aux données des populations d'oiseaux recueillies par l'association ornithologique Aves dans le cadre du programme de Suivi des Oiseaux Communs en Wallonie (SOCWAL). Ce travail s'inscrit dans le cadre du projet "Atlas européen", dont Aves fournit les données concernant la Région wallonne. L'objectif est d'obtenir une cartographie de la densité des populations d'oiseaux à une échelle plus fine que celle publiée par l'association.

2 État de l’art

2.1 Les indicateurs de biodiversité

Un indicateur de biodiversité est un outil aidant à quantifier la biodiversité et à en évaluer les tendances ainsi que les changements au sein d’une communauté ou des écosystèmes afin de les communiquer aux politiques et aux citoyens (GREGORY ET AL., 2005). Le terme biodiversité provient de la contraction de ”diversité biologique”, terme englobant les variations biotiques depuis le niveau des gènes, jusqu’à celui des écosystèmes, en passant par celui des espèces (PURVIS & HECTOR, 2000).

Le terme ”biodiversité” peut être vu comme étant la répartition du nombre d’espèces par surface de territoire. Ainsi, il est plus facile de pouvoir la quantifier. En effet, rapporter ce nombre d’espèces sur une surface peut être une manière de quantifier la biodiversité sur un site, dans un premier temps, et de la comparer avec un autre site ensuite. Cependant, cette méthode fait l’hypothèse grossière que chaque espèce contribue de manière équivalente à la biodiversité de son site (JOHN L. HARPER & DAVID L. HAWKSWORTH, 1994).

Dès lors, il est utile de définir les critères qui définissent un indicateur idéal. Selon CARO et O’DOHERTY (1999), un bon indicateur est un groupe d’espèces dont la tendance des populations :

1. Reflète le comportement moyen des espèces constitutives
2. Met en lumière les tendances des attributs d’autres taxons
3. Agit comme un substitut de la santé de l’écosystème

De manière plus détaillée, le Tableau 1 reprend tous les critères associés à un indicateur idéal de biodiversité.

L’utilisation de ces indicateurs se heurte à certaines limites. Les tendances observées chez une espèce peuvent différer chez une autre espèce pour une variété de raisons ; par exemple, si ces espèces se perpétuent dans des régions différentes, changeant les conditions dans lesquelles elles vivent (KUNIN & GASTON, 1993) ou lorsqu’elles sont sujettes à des mesures de conservation spécifiques (AEBISCHER et al., 2000). En outre, ces indicateurs ne peuvent ni expliquer les causes de ces tendances, ni permettre de formuler des actions de gestion pertinentes. Pour cela, des recherches supplémentaires doivent être entreprises afin d’approfondir les connaissances sur les milieux concernés (GREGORY & STRIEN, 2010).

TABLEAU 1 – Critères d'un bon indicateur selon GREGORY ET AL. (2005).

Attribut	Détails
Représentatif	Inclut toutes les espèces d'un taxon choisi ou un groupe représentatif.
Immédiat	Capable d'être mise à jour régulièrement, idéalement au moins une fois par an.
Information simplifiée	Transparent, facile à interpréter et visuellement attrayant. Les informations complexes doivent être présentées simplement pour avoir un impact et communiquer.
Facilement compréhensible	Les non-experts, qu'il s'agisse de décideurs ou de membres du public, doivent être en mesure de saisir les enjeux pour se les approprier.
Quantitatif	Donne une mesure précise avec évaluation de l'erreur. Montre les tendances dans le temps, mesure un taux de changement et les changements dans ce taux.
Sensible aux changements	Sensible aux changements environnementaux sur des échelles de temps relativement courtes.
À faible délai	Permet une identification rapide des tendances et une alerte précoce des problèmes.
Susceptible d'être analysé	Les données peuvent être désagrégées pour aider à comprendre les schémas sous-jacents et mettre en lumière les causes potentielles des tendances.
Réaliste à collecter	Les données quantitatives sont disponibles ou peuvent être collectées facilement ne nécessitant pas de ressources financières excessives ou irréalistes.
Indicatif	Représentant des composantes ou des attributs plus généraux de la biodiversité que les seules tendances des espèces constitutives, reflétant idéalement la santé des écosystèmes.
Axé sur l'utilisateur	Développés en réponse aux besoins des parties prenantes.
Pertinent sur le plan politique	Les indicateurs visent à fournir des signaux aux clients des politiques pour les aider à élaborer puis à réviser les mesures politiques.
Stable	Protégés des fluctuations naturelles irrégulières et importantes.
Traçable	Sensible à l'influence et au changement humains.

2.2 Les oiseaux comme indicateurs de biodiversité

Afin de garantir une bonne diversité biologique, il faut s'assurer que l'avifaune respecte les critères d'un bon indicateur. L'organisation BirdLife International reconnaît que les oiseaux sont de bons indicateurs d'état de la biodiversité pour les raisons suivantes (BIRDLIFE, 2013) :

1. La taxonomie des oiseaux est bien connue et stable ;
2. Leur distribution, leur écologie et leur cycle biologique sont bien compris ;

3. Ils sont généralement faciles à identifier, à étudier et à surveiller, un ensemble de données historiques précieuses est répertorié pour un large éventail d'espèces ;
4. Ils occupent des territoires très diversifiés, avec des exigences spécifiques, partout dans le monde ;
5. Ils occupent généralement des niveaux trophiques élevés et, par conséquent, sont relativement sensibles aux changements environnementaux ;
6. Les tendances de leurs populations reflètent correctement les tendances d'autres espèces (mammifères, reptiles, plantes, etc.) ;
7. Ils ont une valeur économique importante ainsi qu'une valeur symbolique populaire.

2.2.1 Oiseaux et territoires diversifiés

Comme mentionné précédemment, les oiseaux sont considérés comme de bons indicateurs de biodiversité car ils occupent des territoires très diversifiés présentant des exigences spécifiques. En effet, la distribution et l'abondance des espèces sont directement liées aux limitations physiques de l'habitat. De plus, cet habitat est caractérisé par un ensemble de facteurs environnementaux dont une espèce dépend pour réaliser son cycle de vie (BLOCK & BRENNAN, 1993). Ainsi, l'habitat doit lui fournir des ressources en nourriture suffisantes, des abris de protection contre les intempéries et les prédateurs potentiels, des sites de nidification et des postes de chants. La compétition et l'attraction intraspécifique sont d'autres critères ne relevant pas directement de l'habitat mais qui influencent la sélection de celui-ci par l'oiseau (JONES, 2001).

Un habitat donné offre des ressources nutritives et des sites de nidification communs à plusieurs espèces. Malgré tout, les espèces inféodées à un même habitat peuvent occuper des niches écologiques différentes, disposant ainsi de stratégies distinctes pour les occuper. BLOCK et BRENNAN (1993) prennent comme exemple dans leur article la cohabitation des différentes espèces insectivores dans les forêts feuillues d'Europe, notamment : (1) le Grimpereau des bois qui poursuit les insectes de bas en haut d'un tronc, (2) le Pic épeiche qui tambourine le tronc pour en extraire les larves et (3) la Grive musicienne qui, quant à elle, prend sa nourriture à même le sol dans la litière forestière.

2.2.2 Espèces spécialistes et généralistes

Qu'ils soient migrateurs ou sédentaires, les oiseaux ont la capacité de sélectionner l'habitat qui leur convient le mieux grâce à leur mobilité. En effet, cette dernière leur permet de prospecter un large éventail d'habitats différents. En général, une espèce spécialiste nichera principalement dans un habitat homogène alors qu'une espèce généraliste occupera un environnement plus hétérogène (KASSEN, 2002). Compte tenu de l'homogénéité de son territoire,

une espèce spécialiste sera nettement plus sensible aux perturbations externes qu'une espèce généraliste. A contrario, les espèces généralistes n'ont pas d'exigences particulières en terme d'habitat et peuvent ainsi coloniser tout type de lieu. Moins dépendantes de leur site de niche, elles peuvent peupler des niches écologiques délaissées par les espèces spécialistes (JULLIARD et al., 2006).

2.3 Les méthodes de suivis pour la modélisation des oiseaux

Le suivi de populations d'oiseaux peut se réaliser via deux méthodes différentes : les méthodes se reposant sur le marquage des individus et les méthodes suivant des individus non marqués.

La première méthode consiste à capturer des oiseaux et de les marquer pour ensuite évaluer la dynamique des populations sur la base d'un taux de recapture. Cette méthode, également appelée le "bagueage des oiseaux", est une méthode très connue. L'article de BAILLIE et SCHAUB (2009) fait le point sur ce type de suivi et l'utilisation de ces données. Deux approches sont utilisées, l'une basée sur les récupérations d'oiseaux retrouvés morts (récupérations de morts) et l'autre sur les recaptures vivantes ou des repérages (marquage-recapture). Les bagues ainsi récupérées sur les oiseaux (morts et vivants) sont utilisées pour modéliser les fonctions des taux de survie annuels et des taux de recapture. La difficulté de cette approche réside dans le fait qu'il n'est pas aisé de faire la distinction entre la mortalité et l'émigration permanente au sein de la zone d'étude. De plus, le nombre de captures entraîne systématiquement une sous-estimation du taux de survie, mais ce biais peut être facilement diminué en augmentant la surface d'étude (MARSHALL et al., 2004).

Une fois les données collectées, il est nécessaire d'intégrer au modèle les variables environnementales afin de comprendre les causes à l'origine des modifications des densités des populations. À cela s'ajoute également les variables sur l'abondance de prédateurs, d'agents pathogènes et de concurrents de la même espèce ou d'espèces différentes qui tendent à réguler la population autour du nombre d'individus que l'environnement peut supporter (IAN NEWTON, 1998). Les paramètres de chaque variable peuvent être estimés en maximisant les vraisemblances conjointes résultantes à l'aide d'un filtre de Kalman (BESBEAS et al., 2002). Ce filtre estime les états d'un système dynamique à partir d'une série de mesures incomplètes ou bruitées. Le schéma explicatif (Annexe A) reprend le principe de fonctionnement de ce filtre.

La seconde méthode consiste à effectuer sur le terrain des comptages d'individus détectés. Ce type de monitoring (ne demandant pas de capture) est le plus utilisé car il est peu chronophage et ne nécessite pas de matériel coûteux. Trois versions existent : les points d'écoute (où l'observateur reste statique et enregistre tous les individus vus et entendus), les transects

(où l’observateur suit un itinéraire au sein d’une surface prédéfinie en encodant toutes les espèces détectées) et les méthodes cartographiques (où l’observateur géolocalise directement les individus, ce qui est la méthode employée dans notre étude). Il est également possible de combiner certaines méthodes sur une même période de prise de mesures (SUTHERLAND et al., 2004).

Différents programmes de suivi ornithologique, notamment le programme SOCWAL (cfr Section 2.4), utilisent la méthode des points d’écoute depuis 1990. Le principe général d’une chaîne de points d’écoute, telle que décrite par le programme SOCWAL, consiste en un parcours comportant une quinzaine de points d’arrêts (appelés points d’écoute) où l’observateur note durant cinq minutes toutes les espèces contactées (entendues et/ou vues), leur nombre ainsi que leur comportement. Pour la prise de données sur le terrain, outre une fiche de terrain mise à disposition (Annexe B), des applications mobiles comme *iObs* ou *ObsMapp* (Annexe C) peuvent également être utilisées par les observateurs (AVES, 2015). Les points d’écoute doivent être séparés d’au moins 250 m en forêt et 400 m en zone ouverte afin d’éviter au maximum les doubles comptages (AVES, 2010).

En Région wallonne, les relevés sont organisés deux fois par an, dans des conditions semblables de date, d’heure (en matinée) et de météo (beau temps), suivant une chaîne de 311 points d’écoutes (Annexe D). Afin de faire une analyse fine des tendances des populations, la méthode TRIM (pour *Trends and Indices Monitoring data*), ainsi que les différentes étapes préalables permettant de préparer les données pour l’analyse (récolte des données, mise en forme des fichiers), a été développée pour l’analyse des suivis temporels d’effectifs. Cette méthode permet de prendre en compte d’éventuelles données manquantes. En effet, une partie des points d’écoute peut ne pas avoir été suivie chaque année (PANNEKOEK J. & VAN STRIEN ARCO J., 2005). Grâce à tous ces suivis, la tendance significative à long terme des populations de plus de 80 espèces d’oiseaux nicheurs en Wallonie a pu être modélisée. Les résultats montrent un déclin majeur pour les espèces spécialistes des milieux agricoles et des milieux forestiers (Tableau 2) (ANTOINE DEROUAUX, 2021).

TABLEAU 2 – Nombre d’espèces en fonction des tendances de leur population wallonne, par habitats (ANTOINE DEROUAUX, 2021).

Habitat	Déclin	Stable	Augmentation	Total
Milieux agricoles	12 (70,6%)	2 (11,8 %)	3 (17,6%)	17
Forêt	15 (62,5%)	7 (29,2%)	2 (8,3%)	24
Autres habitats	22 (45,8%)	13 (27,1%)	13 (27,1%)	48
Total	49 (55,1%)	22 (24,7%)	18 (20,2 %)	89

Le programme de surveillance des oiseaux nicheurs en Angleterre utilise la méthode du transect linéaire (SUTHERLAND et al., 2004). La méthode consiste à suivre un itinéraire prédéfini et de relever toutes les espèces observées au sein d’une zone d’étude prédéfinie. En Angleterre, les observations sont inscrites au sein de 3 catégories de distance : 0-25, 25-100 et au-delà de 100 mètres, mesurée perpendiculairement au chemin emprunté. Deux à quatre visites par itinéraires sont recommandées (GREGORY, 2000). Enregistrer les espèces dans des catégories de distance permet une mesure de la présence des oiseaux au sein d’habitats différents et rend l’estimation de la densité de population possible (BIBBY et al., 2000).

L’analyse des changements annuels dans les populations est basée sur le modèle TRIM utilisant le logarithme linéaire de la régression de Poisson. Les résultats montrent un réel déclin dans les populations des contrées anglaises (GREGORY, 2000). L’un des principaux problèmes rencontrés par cette méthode de relevés est que les observateurs ne détectent pas toujours les individus avec précision (y compris dans les milieux ouverts), ce qui peut biaiser l’effectif de population estimé. SANZ-PÉREZ et al. (2020) illustrent bien ce phénomène en comparant le modèle TRIM [qui se base sur l’index annuel des populations tiré du dénombrement des oiseaux] et le modèle HDS (pour *Hierarchical Distance Sampling*) [qui se base sur les estimations d’abondance corrigées pour les détections imparfaites et variables suivant une distribution de Poisson]. Les résultats montrent que, sur 30 espèces d’oiseaux inféodés au milieu agricoles, 43 % des espèces montrent une tendance différente en fonction du modèle utilisé. Cela implique que la détectabilité n’est pas constante lors de recensements. En effet, l’observateur ainsi que l’année du suivi affectent particulièrement la détection et ces effets varient d’une espèce à l’autre.

Ceci démontre l’importance primordiale de tenir compte de la détection imparfaite dans les programmes de surveillance des oiseaux pour garantir des estimations fiables des tendances.

2.4 Les programmes de suivis d’oiseaux

Étant donné l’importance que joue les oiseaux en tant qu’indicateur de biodiversité, différents programmes de suivis ont débuté partout en Europe. En Belgique, chaque Région a son propre programme : l’Algemene Broedvogels Vlaanderen (ABV) en Région flamande ; la surveillance des oiseaux communs de Bruxelles (SOCBRU) en Région Bruxelloise et le programme du suivi des oiseaux communs en Wallonie (SOCWAL) en Région wallonne. Ces trois programmes s’intègrent dans un ambitieux programme international : le Pan-European Common Bird Monitoring Scheme. Ce grand projet de l’European Bird Census Council collecte les relevées de 28 pays européens et calcule les tendances des populations à l’échelle européenne (“PECBMS”, 2022).

En Wallonie, le SOCWAL a débuté en 1990 et est coordonné par Aves, le pôle ornithologique de Natagora, dans le cadre d’une convention avec le Service Public de Wallonie (SPW-DGO3-DEMNA). Dès lors, chaque année, des observateurs suivent une méthodologie spécifique afin de collecter un maximum de données sur l’avifaune wallonne. les tendances annuelles de l’évolution des populations des oiseaux nicheurs peuvent ainsi être calculées. Ces inventaires contribuent à la recherche scientifique et les résultats de ces suivis sont utilisés dans de nombreux travaux sur l’état de la nature : Liste Rouge des oiseaux en Wallonie, Rapport Planète Vivante du WWF, Rapportage pour l’Article 12 de la Directive Oiseaux, État de l’Environnement Wallon, plan d’action européen pour la Tourterelle des bois...) (ANTOINE DEROUAUX, 2021).

2.5 Aves, pôle ornithologique de Natagora

Natagora est une association sans but lucratif active dans tout l’espace Wallonie-Bruxelles qui a pour but de protéger la nature. L’association a été créée le 21 juin 2003 grâce à la fusion de deux associations : (1) Réserves Naturelles RNOB (Réserves naturelles et ornithologiques de Belgique) qui se consacraient à la préservation et à la gestion des habitats naturels menacés, et (2) Aves qui organisait des excursions dans la nature à destination du public bruxellois. Le développement de cette association a pour but de reprendre toute une série de missions, autrefois assurées par les deux associations mères (“Réserves Naturelles RNOB / Natagora (organisme) - Réseau IDée asbl”, s. d.).

L’objectif premier de Natagora est d’enrayer la dégradation de la biodiversité et reconstituer un bon état général de la nature, en équilibre avec les activités humaines. Par ses actions de protection et de sensibilisation, Natagora crée et gère des réserves naturelles ayant reçu le statut juridique le plus strict et le plus durable prévu par la législation. Plus de 5000 hectares de réserves abritant une biodiversité très diversifiée sont gérés au quotidien par des professionnels et volontaires. Ces réserves sont managées au cas par cas en fonction du milieu et des espèces présentes sur le site. Les tâches réalisées sont multiples et diversifiées : inventaires, préparation et mise en œuvre de travaux de gestion, développement de partenariats, gestion par pâturage, suivi biologique de l’évolution des milieux, surveillance, etc. Plusieurs autres actions impliquant tous les acteurs de la société sont développées, tels que des événements ponctuels, des groupes de travail thématiques, des campagnes ciblées ou des initiatives pour allier les arts contemporains et la science (“Natagora”, 2018).

Aves est une association fondatrice de Natagora dont elle constitue le pôle ornithologique. Née en 1953, elle organisait dans un premier temps des excursions dans la nature à destination

du public. Dès 1956, Aves revendique son envie de protéger l’avifaune. Dans cette optique, sa première mission concrète a été de protéger les faucons pèlerins (*Falco peregrinus*) contre les collectionneurs d’œufs et certains fauconniers. Des banques de données ornithologiques ont d’emblée été tenues par des bénévoles afin de publier un bulletin contenant des résultats de recensements, des enquêtes ou des articles sur des événements marquants (passage de certaines espèces rares par exemple). Vers la fin des années 90, les données passèrent en format numérique, ce qui permit en 2011 de publier l’*Atlas des oiseaux nicheurs de Wallonie*, commandé par la Région wallonne avec qui Aves collabore depuis 1996 pour la surveillance des indicateurs de la biodiversité en Wallonie (programme SURWAL¹). De nos jours, le monitoring des populations d’oiseaux en Wallonie et à Bruxelles constitue l’une des principales activités d’Aves (“Aves, pôle ornithologique de Natagora, association de protection de la nature”, s. d.).

En partenariat avec de nombreuses collaborations nationales et internationales (LifeWatch², BirdLife, Natuurpunt, DEMNA, etc.) pour des études, mais également pour la protection des espèces, les données recueillies par Aves dans le cadre du programme de Suivi des Oiseaux Communs en Wallonie (SOCWAL) permettent de calculer des indicateurs de l’état de la biodiversité (DEROUAUX & PAQUET, 2018).

2.5.1 Les relevés de données oiseaux

Au cours des recherches de terrain de l’Atlas des Oiseaux Nicheurs en Wallonie 2001-2007, des dizaines d’observateurs professionnels ont réalisé deux relevés standardisés d’une heure chacun. Le premier passage devait avoir lieu avant le 30 avril et le deuxième après le 15 mai à l’intérieur de carrés (2800 représentant 17 % du territoire wallon) de 1x1 km, répartis de manière régulière dans toute la Wallonie (Annexe E). Durant leur passage, les observateurs, suivant les règles d’un protocole strict, ont noté tous les individus vus ou entendus au sein de ces carrés. De cet effort collectif a résulté la sortie de cartes à haute résolution (Figure 2) présentant la densité relative des espèces répandues. Ces données permettent de mieux comprendre la connexion entre les oiseaux et leur habitat. De plus, elles constituent une base extraordinaire pour comprendre l’évolution récente de l’avifaune en région wallonne.

À partir de 2015, un nouvel échantillonnage a été mis en place. Les nouveaux relevés ont été menés dans les conditions les plus similaires possibles par rapport à ceux menés pendant les années 2001-2007. Deux relevés sont toujours menés entre le 25 mars et le 30 avril (10

1. Activités de monitoring et de surveillance de la biodiversité entamé par l’Etat de l’Environnement Wallon fin des années 90 (“La biodiversité en Wallonie”, s. d.).

2. cfr Section 2.6

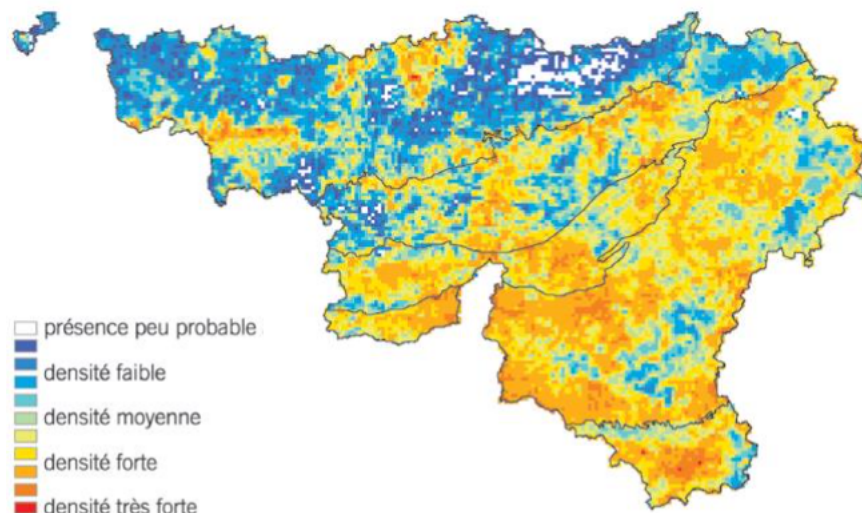


FIGURE 2 – "Carte présentant la densité relative du Rouge-gorge familier (*Erithacus rubecula*) en Wallonie en 2001-2007, réalisée grâce aux échantillonnages kilométriques. Cette espèce est présente partout, mais montre de fortes variations de densité au travers du territoire, essentiellement liées à la couverture forestière" (AVES, 2015).

mai en Haute Belgique) pour le premier, entre le 15 mai et le 30 juin pour le second d'une heure chacun. Cependant, ce ne sont plus 2800 carrés qui sont observés, mais 1600. Ces 1600 nouveaux carrés représentent un peu moins 10 % du territoire wallon. Selon un travail de recherche mené par AIZPURUA et al. (2015) sur les données du premier échantillonnage de l'atlas wallon, il s'avère qu'une couverture de l'ordre de 4-5 % (comparée au 17 % obtenu entre 2001-2007) est déjà suffisante pour modéliser la répartition des espèces peu ou moyennement répandues. Autrement dit, il n'est pas nécessaire de réitérer l'ensemble de l'échantillonnage ce qui permet de gagner du temps (AVES, 2015).

2.6 Les projets LifeWatch

LifeWatch ERIC (pour *European Research Infrastructure Consortium*) est un consortium européen d'infrastructures de recherche établi par la Commission européenne en 2017. Il permet de fournir des installations de recherche électroniques aux scientifiques étudiant la biodiversité et les fonctions et services écosystémiques afin d'aider la société à relever les principaux défis planétaires (LIFEWATCH, s. d.).

Malgré son territoire relativement petit, la Belgique possède une remarquable diversité d'habitats et d'espèces qu'elle souhaite protéger efficacement, tant à l'intérieur qu'à l'extérieur de ses frontières. Depuis le lancement des projets LifeWatch, la Belgique s'est montrée engagée, soutenue par les autorités politiques respectives, et contribue activement à la recherche à

travers un certain nombre de projets à long terme gérés par différents centres de recherche et universitaires. LifeWatch Belgium est même devenu l'un des sept nœuds nationaux composant LifeWatch ERIC avec les contributions de la région Flamande, de la Fédération Wallonie-Bruxelles et de l'État fédéral. Les cinq grands axes de recherche développés et exploités par le centre national LifeWatch en Belgique sont (LIFEWATCH, s. d.) :

1. Un réseau d'informations sophistiqué sur les espèces ;
2. Un observatoire des eaux douces et marines ;
3. Une installation de cartographie de l'habitat sur base de télédétection ;
4. Un laboratoire virtuel d'observation de la biodiversité en Antarctique ;
5. Une installation d'identifications d'organismes ou de tissus dérivés et de produits d'intérêt politique.

Afin de préserver cette diversité, les plans de gestion et de conservation se basent sur des modèles statistiques. Depuis 2012, dans le but de toujours améliorer la précision et la justesse de ces modèles, la Fédération Wallonie Bruxelles contribue en nature³ à l'infrastructure de recherche Lifewatch. Grâce à ces recherches, la capacité prédictive des modèles s'est en effet améliorée. Cependant, ceux-ci ont besoin d'informations sous-jacentes fiables pour fonctionner correctement. Afin que ces informations soient cohérentes et de qualité, un nouveau type de base de données géographiques a été conçu par LifeWatch en lien avec le concept d'écotopie (WALLONIE-BRUXELLES, 2019).

2.6.1 Les écotopes

Dans le processus de création et de diffusion des cartes thématiques, l'évaluation de la précision est une condition nécessaire à la comparaison des résultats de la recherche et à l'utilisation appropriée des produits cartographiques (STEHRMAN & WICKHAM, 2011). Dans ce processus d'évaluation, trois étapes importantes mènent à la conception d'une base de données géographiques (PETER A. BURROUGH & A. FRANK, 1996) : (i) l'observation, (ii) la mise en relation de l'observation avec un modèle conceptuel, par exemple GEOBIA (pour *Geographic Object-Based Image Analysis*) (CASTILLA et al., 2012), et (iii) la représentation des données en termes formels.

Lors de la première étape, l'observation doit être identifiée en tant qu'entité clairement définie (entités spatiales) ou définissable (régions spatiales ou continuum). Le continuum, surface

3. "Les contributions volontaires en nature (CVN) représentent l'acte par lequel une personne physique ou morale apporte à l'association des biens ou d'autres services à titre gratuit. Elles sont donc attribuées en considération du statut désintéressé de l'organisme. Ces apports doivent être mis en valeur dans les comptes annuels des associations" ("Les contributions volontaires en nature dans les associations", 2018).

variant faiblement, est appelé ici champ continu. Cependant, voir le monde comme un continuum ou comme un ensemble d'entités discrètes est toujours une question d'appréciation (COUCLELIS, 1996), où certaines frontières peuvent difficilement être délimitées, comme le cas de la couverture et de l'occupation du sol. En effet, ces catégories peuvent être discrétisées comme des polygones vectoriels avec des frontières basées sur un accord général (qui sont représentées comme des lignes adimensionnelles, mais qui sont floues en réalité), ou elles peuvent être représentées comme une grille arbitraire spécifiant la proportion de chaque entité individuelle à chaque pixel (sans limites) (BIAN, 2007). Ces trois modèles conceptuels (entité spatiale, région spatiale et champ continu) sont illustrés à la Figure 3.

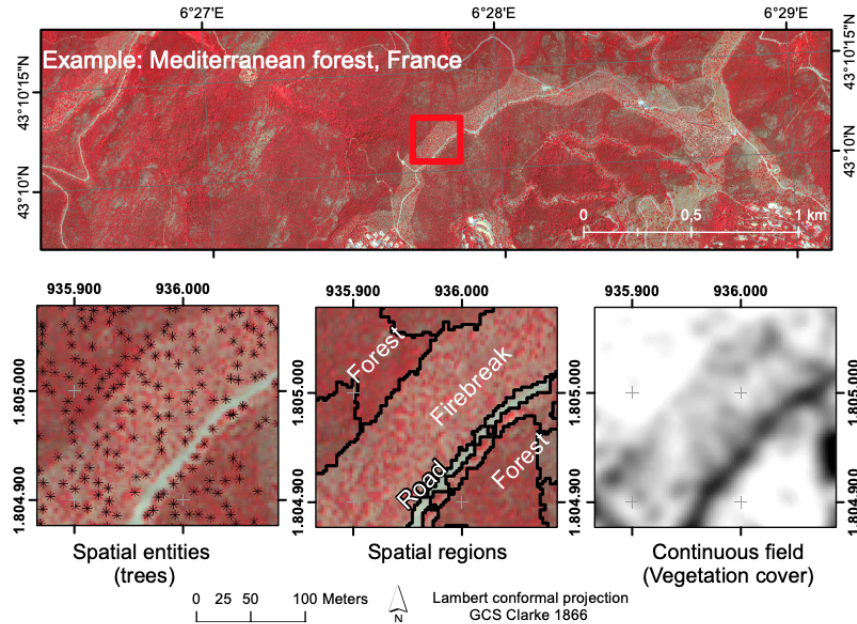


FIGURE 3 – Illustration des modèles conceptuels d'une forêt méditerranéenne dans le sud de la France. Dans ce paysage couvert par une image Ikonos, les arbres sont des entités spatiales détectables individuellement (étoiles sur l'image de gauche). Au centre, la segmentation automatisée des images crée des régions spatiales qui correspondent à l'ancien pare-feu et à la forêt non perturbée composée d'arbres, d'arbustes et de végétation herbacée. À droite, une représentation continue du champ met en évidence la densité de végétation plus faible (gris plus foncé) dans l'ancien pare-feu (JULIEN RADOUX & PATRICK BOGAERT, 2017).

Relatif au modèle conceptuel des géo-objets, les entités spatiales ou les régions spatiales sont représentés (et analysés) sous forme de points, de lignes ou, la plupart du temps, de polygones. Ensuite, un ensemble unique d'attributs est adjoint à chaque polygone dans son ensemble, contrairement à la classification basée sur les pixels, où chaque pixel individuel est classé. Les polygones sont donc intrinsèquement évalués comme homogènes en termes de label (SCHÖPFER et al., 2008), sont une bonne représentation des entités spatiales et

sont une représentation raisonnable pour les régions spatiales, qui peuvent avoir des limites indéterminées (BIAN, 2007). Particulièrement pour ces régions spatiales, diverses classes (par exemple, zone urbaine, forêt ouverte, forêt mixte, verger) sont également définies au niveau du polygone. Cette classification comprend généralement des concepts complexes où le contexte joue un rôle majeur pour l'interprétation du résultat (DURO et al., 2012).

Dans ce mémoire, les entités spatiales sont spécifiées comme écotopes. Les écotopes⁴, unités territoriales, sont identifiables sur le terrain en fonction de la couverture du sol, de l'utilisation du sol, de la topographie et des types de sols. Leur modélisation écologique a permis de déterminer leur délimitation et leur caractérisation. Cette délimitation est importante. En effet, il a été démontré que la taille et que la forme des objets (polygones) ont leur importance dans la prédiction d'abondance des espèces (DELANGRE et al., 2018). Les écotopes sont structurés en polygones, méthode bien adaptée à la modélisation des habitats, auxquels sont associés différents descripteurs quantitatifs de l'occupation et de l'utilisation du sol. Ces polygones ont été délimités automatiquement sur base du relief LIDAR et des orthophotographies. Comme le montre la Figure 4, à chaque polygone est associé une base de données reprenant les différentes variables quantitatives issues des nombreuses couches d'information disponible sur le territoire (occupation du sol, pente, drainage, temps d'ensoleillement, etc.) (WALLONIE-BRUXELLES, 2019). La classification des écotopes présents dans la zone d'étude est disponibles à l'Annexe F.

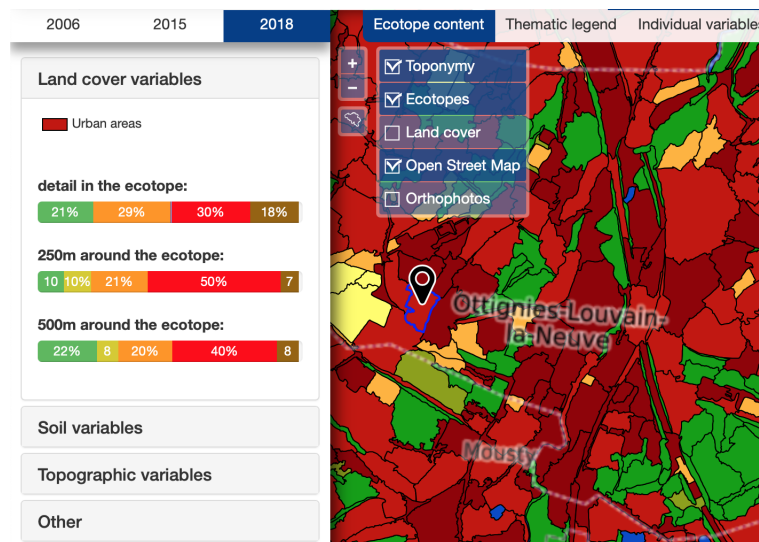


FIGURE 4 – Affichage des attributs d'un écotope avec les proportions de classes d'occupation du sol.

4. "Les écotopes sont définis comme étant les plus petites unités du paysage écologiquement homogènes" (RADOUX et al., 2019)

2.7 Prédire l'abondance des espèces

Les processus biologiques se déroulent sur des périodes allant de la microseconde à des millénaires et sur des échelles allant du nanomètre au continent. Cependant, la plupart des processus écologiques se produisent sur une gamme plus limitée d'échelles spatiales et temporelles qui dépendent des processus et des organismes spécifiques impliqués. Par conséquent, la compréhension de la dynamique et de la régulation d'un processus tel que la distribution des espèces nécessite des informations précises concernant l'occupation du sol et les ressources naturelles disponibles, mais également des informations sur les interactions entre espèces et les conditions environnementales (J. MICHAEL SCOTT et al., 2001). Ces facteurs sont incorporés dans des modèles explorant les relations entre l'occurrence des espèces et l'ensemble des variables prédictives de ces processus. Deux types de résultats peuvent être tirés de ces modèles : (1) la probabilité d'apercevoir une espèce dans un lieu non exploré par extrapolation des données ; (2) des estimations de la pertinence d'une zone pour les espèces (SEGURADO & ARAÚJO, 2004).

Les données de surveillance à long terme de la biodiversité sont principalement utilisées pour estimer les changements dans l'occupation ou l'abondance dans l'espace des espèces au fil du temps (AIZPURUA et al., 2015). Elles ont une importance d'autant plus cruciale qu'elles peuvent prévenir du déclin et/ou de l'extinction imminents d'espèces, permettant ainsi de déclencher des interventions de gestion (LINDENMAYER et al., 2012). Ces données peuvent également être intégrées dans des modèles prédictifs pour étudier la répartition des espèces dans l'espace. Bien que les changements d'occupation ou d'abondance puissent être estimés à partir d'un nombre relativement limité d'unités d'échantillonnage, la petite taille de l'échantillon peut conduire à des modèles spatiaux erronés et des cartes de distribution d'espèces observées inexactes (AIZPURUA et al., 2015).

Ces cartes des distributions permettent un suivi scientifique dans le but de protéger la nature avec pour objectif d'enrayer la dégradation de la biodiversité et de reconstituer un bon état général de la nature, en équilibre avec les activités humaines ("Natagora", 2018). En effet, ce suivi à long terme de la faune sauvage est généralement considéré comme un outil essentiel pour la gestion de la biodiversité et pour les études de recherche sur la conservation de la biodiversité (LINDENMAYER et al., 2012).

2.7.1 Aperçu des modèles Espèces-Habitats

Les études espèces-habitats examinent généralement les caractéristiques environnementales associées à la distribution d'une espèce donnée. Elles cherchent avant tout à prédire la distribution actuelle des espèces, ensuite à la projeter selon des scénarii de changement global.

Elles permettent alors de démontrer les conséquences sur les écosystèmes ou les modifications au sein des populations liées à des perturbations naturelles ou anthropiques (HATTAB, T. et al., 2011).

Par exemple, BALDERAMA et al. (2016) ont étudié la distribution, l'abondance et la variabilité spatio-temporelle des oiseaux marins dans leurs habitats hauturiers liées à l'influence d'installations d'énergie éolienne offshore au large des États-Unis. Ils en ont conclu qu'il est de plus en plus nécessaire de caractériser précisément ces paramètres démographiques avant d'évaluer l'influence de facteurs tels que le développement énergétique offshore sur les populations aviaires. GRINDE et al. (2017) ont démontré l'interaction significative entre les conditions météorologiques et la couverture terrestre sur le recensement des oiseaux. Dans l'ensemble, leurs modèles ont indiqué que les variables de la couverture terrestre et de la biomasse aérienne des arbres entraînaient des variations pour la majorité des espèces. GUIBAN et al. (2002) ont tenté de définir la gamme d'applications (dans l'espace et dans le temps) pour lesquelles les prédictions du modèle sont adaptées. Selon eux, une précision et une résolution plus élevées des cartes d'entrées biophysiques sont toujours considérées comme les principales exigences pour améliorer les prévisions des modèles.

Concernant les facteurs naturels, il semble que le changement climatique et le changement d'affectation des terres sont les deux principaux facteurs de perte de biodiversité (HABOU et al., 2021). Plus précisément, le changement climatique a le potentiel de modifier la structure et le fonctionnement d'un écosystème, en ayant des impacts à l'échelle mondiale et régionale importants sur la biodiversité. Il s'avère également que les impacts, suite à une modification de la couverture du sol et des perturbations météorologiques, sont spécifiques et varient en fonction des espèces. Dans l'ensemble, les modèles ont démontré que les variables de la biomasse aérienne et de la couverture terrestre entraînaient des changements chez la majorité des espèces. Quant à la variabilité annuelle des conditions météorologiques, celle-ci est beaucoup plus susceptible d'avoir un impact sur l'abondance des espèces associées aux forêts de feuillus et aux espèces d'oiseaux considérées comme sensibles au climat (GRINDE et al., 2017). De plus, ces études peuvent mener à des résultats différents en fonction des paramètres environnementaux sélectionnés à différentes échelles ainsi que la durée de l'expérience. En effet, lors d'une expérimentation en Estonie, une étude a pu le démontrer en comparant l'impact de la fragmentation des milieux sur des espèces d'oiseaux. Il en ressort que les probabilités qu'il y ait des effets de fragmentation sur un succès de nidification sont différents en fonction de l'échelle à laquelle cette fragmentation a été testée ainsi que le temps d'étude (OJA et al., 2005). À haute résolution, des données telles que des arbres ou des haies peuvent améliorer l'analyse des relations espèces-habitat et rendent potentiellement les cartographies

plus précises, alors qu'à une plus faible résolution, certaines caractéristiques environnementales ne sont pas prises en compte, biaisant ainsi les résultats (GOTTSCALK et al., 2011). Il est donc intéressant de travailler à haute résolution afin de comprendre correctement les liens entre espèces et habitats.

2.7.2 L'importance de travailler à haute résolution

Dans son étude, GOTTSCALK et al. (2011) a fait varier la résolution spatiale des pixels de 1 mètre à 1000 mètres dans des modèles espèces-habitat de 13 espèces d'oiseaux nichant dans le centre de l'Allemagne. L'étude a pu démontrer qu'en accroissant la taille de ces pixels, la déviance expliquée⁵ (D^2) des modèles d'oiseaux a généralement diminué suite à une résolution spatiale plus grossière des données (Figure 5). En effet, en augmentant la taille des pixels, le pourcentage de couverture des classes d'affectation du sol communes a augmenté, tandis que celui des affectations des terres plus rares a diminué. Ainsi, 21,8 % de forêts de feuillus était cartographié à 1 mètre de résolution. Ce pourcentage atteint 35,5 % à 1000 mètres de résolution. À l'opposé, les forêts de conifères ont diminué de 6,6 % à 3,5 %.

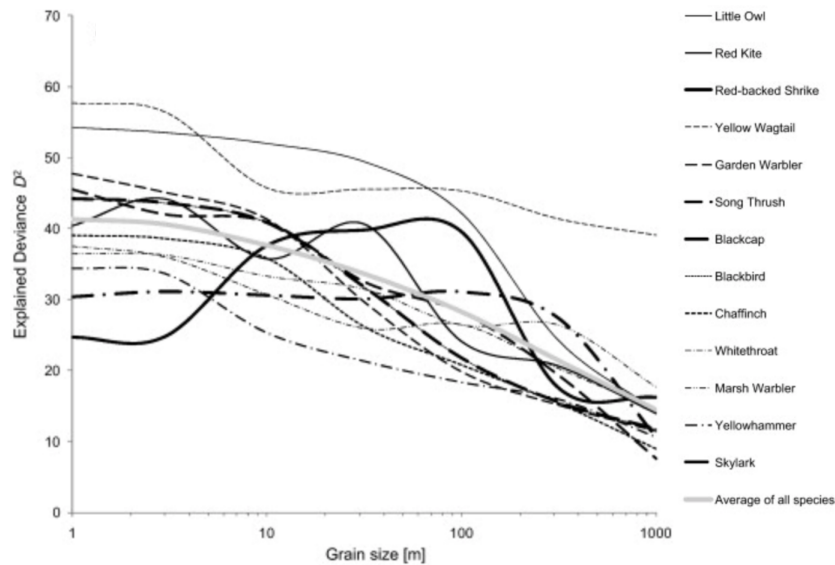


FIGURE 5 – Influence de la taille des pixels sur la déviance expliquée, où chaque courbe correspond à une espèce d'oiseau différente. En moyenne, pour toutes les espèces, la déviance expliquée diminue de 41.3 (à 1 mètre de résolution) à 14.4 (à 1000 mètres de résolution) (GOTTSCALK et al., 2011).

5. Cette fonction calcule la quantité (ajustée) de déviance prise en compte par un modèle linéaire généralisé ; plus sa valeur est faible, au moins le modèle est performant.

Les résultats de l'étude précisent également que 10 espèces étaient correctement modélisées à 1 mètre de résolution, 2 espèces l'étaient à 3 mètres et une seule espèce l'était à 32 mètres de résolution. Au-delà de cette valeur critique, le modèle ne modélisait plus adéquatement aucune espèce d'oiseau.

En revanche, la combinaison de 3 facteurs a sûrement permis une haute performance du modèle à des tailles de pixels plus fins. Ces 3 facteurs sont : (1) une précision spatiale élevée des enregistrements d'oiseaux, (2) un emplacement et une délimitation plus précis des caractéristiques de l'habitat et, (3) dans une moindre mesure, un plus grand nombre de types d'habitats différenciés dans des cartes de résolution plus fine.

Il est donc essentiel de minimiser les tailles des pixels ou des entités spatiales afin de rendre un modèle prédictif performant et d'obtenir des cartographies précises à haute résolution, d'où l'intérêt de ce travail.

2.7.3 Comparaison entre différents modèles statistiques de régression

Les écologistes tentent d'étudier les nombreuses relations entre l'occurrence et/ou l'abondance d'une espèce et son environnement. La majorité de ces études utilise diverses techniques de régression dans un but de développer un modèle statistique valide qui peut être utilisé pour prédire ou interpréter un processus écologique (J. MICHAEL SCOTT et al., 2001).

La complexité et la variation inhérentes aux espèces et à leurs réponses aux facteurs physiques et biologiques à plusieurs échelles rendent difficile la prédiction de l'occurrence, de l'abondance ou de la viabilité des espèces, avec des niveaux élevés de précision et d'exactitude (J. MICHAEL SCOTT et al., 2001).

Une étude menée sur une période de 5 ans a montré que 43 % des modèles de prédiction d'abondance utilisent une régression linéaire multiple, 27 % utilisent une régression logistique, 34 % utilisent d'autres techniques paramétriques, et 2 % utilisent une classification et un arbre de régression. MORRISON et al., 2012 confirment que les techniques statistiques les plus utilisées pour développer des modèles prédictifs relatifs à l'abondance d'une espèce au sein de son habitat sont des régressions linéaires multiples. Plusieurs raisons peuvent expliquer que ces dernières sont les plus courantes (J. MICHAEL SCOTT et al., 2001) :

1. L'utilisation historique et l'acceptation de cette technique dans la littérature scientifique ;
2. La facilité d'utilisation de cette technique en statistique par les écologistes ;
3. La facilité d'interprétation des résultats.

Cependant, il reste difficile à déterminer quel modèle est le plus adapté. Un modèle inapproprié peut entraîner des prédictions erronées sur l’occurrence des espèces dans un milieu. Il est donc important de veiller à sélectionner des modèles ad hoc (LIN et al., 2015). Un panel de modèles prédictifs a donc été développé par les scientifiques afin de comprendre et d’expliquer la corrélation entre les réponses des modèles et les variables prédictives (GUISAN & ZIMMERMANN, 2000).

POTTS et ELITH (2006) ont comparé 5 modèles de régression (PRM pour *Poisson Regression Model*, NBR pour *Negative Binomial Regression*, QPR pour *Quasi-Poisson Regression*, HM pour *Hurdle Model* et ZIP pour *Zero-Inflated Poisson*) afin d’évaluer la relation entre l’abondance d’une espèce végétale vulnérable, *Leionema ralstonii*, petit arbuste aux feuilles vert pâle (“PlantNET - FloraOnline”, s. d.), et son environnement. De cette étude, il est ressorti que le HM donne les meilleurs résultats, avec les corrélations les plus élevées entre les abondances observées et prédites. À l’opposé, le NBR a été le moins performant avec des corrélations plus faibles que les autres modèles et un biais fort. Quant au PRM, il a donné des estimations précises de l’abondance de la population.

De nos jours, des modèles radicalement différents des modèles linéaires généralisés (GLM pour *Generalized Linear Model*) font leur apparition. Les analyses basées sur les arbres de décision décrivent une partition des observations suivant des conditions logiques sur les variables explicatives. Des méthodes élaborées, comme le cas des forêts d’arbres décisionnels (*random forest*), permettent de réduire les variances associées aux estimations du modèle (BARBU et al., 2019). Afin de prédire la distribution d’espèces rares telles que 3 espèces de grues (Grue moine *Grus monacha*, Grue à cou blanc *Antigone vipio* et Grue à cou noir *Grus nigricollis*) en Chine, MI et al. (2017) ont évalué quatre algorithmes d’apprentissage automatique afin de cartographier les distributions de reproduction des trois espèces : TreeNet (*Stochastic Gradient Boosting, Boosted Regression Tree Model*), Random Forest, CART (*Classification and Regression Tree*) and Maxent (*Maximum Entropy Models*). Au final, leurs résultats ont démontré que le Random Forest a la meilleure performance pour la plupart des méthodes d’évaluation puisqu’il a fourni un meilleur ajustement du modèle aux données de tests et a obtenu de meilleures cartes de répartition des espèces pour chaque espèce de grue.

2.7.4 La régression de Poisson et le modèle linéaire généralisé

Grâce au développement technologique, les scientifiques ont davantage de facilité à utiliser et à accéder aux techniques de régression modernes qu’auparavant. Des techniques de régression basées sur des modèles linéaires généralisés comme la régression de Poisson sont de plus en plus utilisées. Ces modèles sont des modèles paramétriques. Effectivement, ils se

basent sur le fait que les données sont conformes à une fréquence de distribution particulière (J. MICHAEL SCOTT et al., 2001). Bien que l'utilisation du modèle de régression de Poisson (PRM pour *Poisson Regression Model*) soit peu commune en écologie (WELSH et al., 1996), celui-ci présente certaines caractéristiques intéressantes qui expliquent la pertinence de son utilisation dans des applications écologiques. En effet, la régression de Poisson est un modèle de prédiction qui s'applique lorsque la variable cible Y est une variable de comptage (plus de détails concernant le modèle, cfr Section 8.1) ("STAT 501 - Regression Methods", s. d.).

De plus, ce modèle s'applique très bien lorsque les données de présence uniquement et les données de comptage sont regroupées sans discernement. Quant au support spatial, la régression de Poisson permet de prédire l'abondance attendue parmi les cellules ou entités spatiales en fonction de leurs covariables spatiales. Dès lors, les estimations des coefficients de ces covariables dépendent de l'échelle spatiale utilisée et l'inférence est uniquement valable pour la résolution spatiale des cellules d'une zone d'étude (FLETCHER et al., 2019).

3 Objectifs du travail

L'objectif principal de ce mémoire est de pouvoir mesurer l'abondance des espèces d'oiseaux nicheurs en Wallonie par une régression de Poisson à une échelle spatiale fine. Cette nouvelle échelle est calibrée sur les écotopes dont la taille moyenne fait 1,35 hectare, soit une diminution d'échelle d'un facteur 100 par rapport au modèle utilisé par Aves dont la résolution spatiale est d'un kilomètre carré.

4 Plan du travail

La méthodologie retenue dans ce mémoire consiste à travailler par écotopes. Les différents écotopes étudiés ont été décrits et classifiés à la Section 2.6.1. Une cartographie de la délimitation et de la classification des écotopes a également été présentée dans cette section.

Concernant la phase de récolte des données, des observations existantes ont été collectées dans les bases de données biologiques, principalement celles de la plateforme Observations.be. Par la suite, des séances de terrain (de mars 2022 à mai 2022) ont été réalisées⁶. Cette campagne de terrain a suivi le protocole officiel instauré par Aves. La nouveauté réside dans le fait que la position géographique de chaque observation est également renseignée. Grâce à cette

6. Il est important de préciser que cette démarche expérimentale se base sur une seule campagne de terrain. Une partie de la discussion est prévue à cet effet.

position géographique, les données sont reliées à l'écotope dans lequel elles ont été aperçues. Pour chaque écotope, des informations concernant l'occupation du sol et d'autres variables environnementales quantitatives sont disponibles. Ces variables permettront de calibrer le modèle prédictif des populations de chaque espèce.

Les résultats seront présentés par milieu présent dans la zone d'étude, à savoir le milieu urbain, le milieu agricole, le milieu forestier et le milieu humide. Ces résultats du nouveau modèle pourront alors être comparés avec les résultats obtenus en 2015 par le modèle employé par Aves. Ainsi, une première validation du nouveau modèle sera effectuée.

Outre la validation des résultats, une analyse de sensibilité permettra d'identifier les potentielles améliorations à apporter au modèle pour des manipulations futures.

Matériels et méthode

5 Sites d'étude : les carrés kilométriques

En Wallonie, la surveillance des oiseaux communs nicheurs est assurée au sein de carrés fictifs dont les dimensions sont respectivement d'un kilomètre sur un kilomètre. Ces unités kilométriques ont été régulièrement réparties pour la période de suivis 2001-2007 sur de grandes diagonales traversant la Wallonie de Nord-Ouest en Sud-Est. Bien qu'il y ait un nouvel échantillonnage introduit en 2015 (cfr Section 2.5.1), les relevés pour ce travail sont entrepris dans les carrés, définis pour la période 2001-2007, de le Nord-Est du Brabant wallon. Une surface carrée de 13x13 kilomètres a été définie (dont les coordonnées en Lambert 72 s'étendent de y : N 165670 ; S 152675 / x : E 167508 ; O 180513) afin de délimiter la zone d'étude. Finalement, 14 carrés au sein de cette superficie ont été visités (en bleu sur la carte) durant les prélèvements s'étendant sur plusieurs communes telles que Grez-Doiceau, Hamme-Mille et Chaumont-Gistoux.

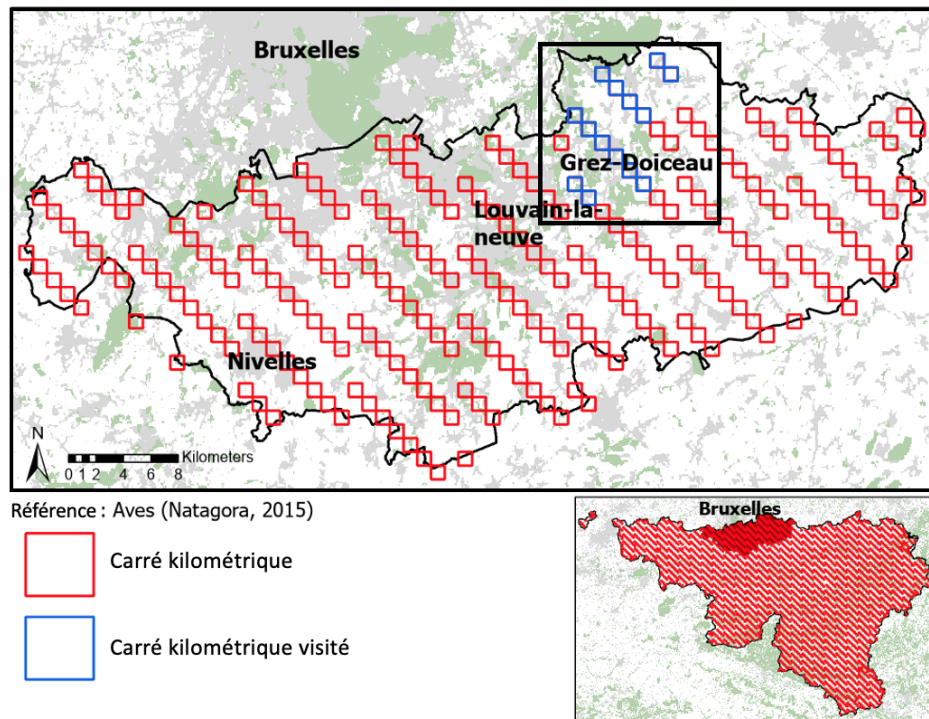


FIGURE 6 – Délimitation de la surface d'étude de 13x13 km (carré noir en surbrillance) et localisation des 14 carrés visités (en bleu) pour l'échantillonnage (AVES, 2015).

6 Choix des espèces observées

Afin de rendre le modèle le plus juste et précis possible, un nombre limité d'espèces est retenu pour le prélèvement des données. La sélection des espèces a été réalisée de manière à avoir un recueil d'espèces varié en fonction de leur milieu.

Les informations relatives aux espèces étudiées, ainsi que les photographies (Annexes /G/H/I/J), proviennent du guide ornithologique (Guide nature, 2019), des fiches des espèces du site “Portail Biodiversité en Wallonie : espèces, biotopes, sites de grand intérêt biologique, actualités” (s. d.) et du site OISEAUX.NET (2022).

6.1 Oiseaux du milieu urbain

TABLEAU 3 – Liste des espèces du milieu urbain.

	Famille	Espèce	Description
Milieu urbain	Passéridés	Moineau domestique (<i>Passer domesticus</i>)	Certainement l'une des espèces les plus anthropophiles. Il s'observe souvent en banc.
		Moineau friquet (<i>Passer montanus</i>)	Plus farouche et lié aux espaces plus ruraux, il se différencie de son congénère par un plumage brun chocolat sur la tête alors que celle du Moineau domestique est grise
	Colombidés	Pigeon biset domestique (<i>Columba livia domestica</i>)	Arpente les milieux bâtis où ils profitent de l'abondante ressource procurée par les activités humaines.
		Pigeon ramier (<i>Columba palumbus</i>)	Idem Pigeon biset domestique
		Tourterelle turque (<i>Streptopelia decaocto</i>)	Idem Pigeon biset domestique
	Muscicapidés	Rougequeue noir (<i>Phoenicurus ochruros</i>)	Le rougequeue noir est un petit passereau au plumage sombre et discret, se trouvant fréquemment en environnement humain.

6.2 Oiseaux du milieu ouvert et agricole

TABLEAU 4 – Liste des espèces du milieu ouvert et agricole.

Milieu ouvert et agricole	Familles	Espèces	Description
	Corvidés	Corneille noire (<i>Corvus corone</i>)	Espèce sédentaire ne passant pas inaperçue. Son cri est puissant et bien caractéristique.
		Choucas des tours (<i>Coloeus monedula</i>)	Le plus petit des corvidés indigènes, oiseau sociable et bruyant.
		Pie bavarde (<i>Pica pica</i>)	Occupe tous les habitats ouverts et semi-ouverts à condition qu'il y ait au moins quelques ligneux pour la nidification.
	Fringillidés	Linotte mélodieuse (<i>Linaria cannabina</i>)	Oiseau commun qui habite toutes sortes de milieux ouverts à semi-ouverts. Son régime alimentaire est exclusivement granivore.
	Emberizidés	Bruant jaune (<i>Emberiza citrinella</i>)	Principalement granivore, est un passereau des campagnes boisées et des lisières. Il chante perché au sommet d'un arbre ou d'un buisson. Son chant un peu mélancolique est facilement reconnaissable.
	Alaudidés	Alouette des champs (<i>Alauda arvensis</i>)	Originaire des steppes, d'où son attrait pour les espaces ouverts. Son chant est très varié pouvant durer plus de 10 minutes sans interruption.
	Sturnidés	Etourneau sansonnet (<i>Sturnus vulgaris</i>)	Son alimentation au sol nécessite des milieux ouverts et des cavités pour sa reproduction. Chanteur et très vocal, son répertoire est large. Il est capable d'imiter le chant d'autres oiseaux, mais également des bruits comme des klaxons de voiture.
	Phasianidés	Faisan de Colchide (<i>Phasianus colchicus</i>)	Très convoité pour la chasse. Il niche caché à même le sol dans les hautes herbes.
	Falconidés	Faucon crécerelle (<i>Falco tinnunculus</i>)	Petit rapace diurne des paysages ouverts. Il chasse en vol stationnaire, principalement des petits mammifères. Il peut compléter son régime d'insectes, de grenouilles et de vers.
	Accipitridés	Epervier d'Europe (<i>Accipiter nisus</i>)	Chasse soit à l'affût en lisière de forêt, soit en repérage aérien des petits passereaux, pigeons ou gros coléoptères. Il surprend ses proies par un vol rasant et rapide, atteignant des vitesses de 70 km/h.
	Accipitridés	Buse variable (<i>Buteo buteo</i>)	Rapace diurne plutôt solitaire nichant dans le milieu forestier (bois, bosquet). Son régime se compose de mammifères de tailles variables, d'amphibiens et de reptiles.

6.3 Oiseaux du milieu forestier

TABLEAU 5 – Liste des espèces du milieu forestier.

Milieu forestier	Famille	Espèces	Description
	Picidés	Pic vert (<i>Picus viridis</i>)	Se différencie grâce à son plumage dorsal vert et le haut de sa tête rouge.
		Pic épeiche (<i>Dendrocopos major</i>)	Plus petit, il a un plumage strié blanc et noir.
	Paridés	Mésange bleue (<i>Cyanistes caeruleus</i>)	La famille des paridés sont des passereaux de taille relativement petite aux couleurs distinctes. Leur régime se compose d'insectes et d'invertébrés pouvant se compléter par des fruits et des graines.
		Mésange charbonnière (<i>Parus major</i>)	Idem Mésange bleue.
		Mésange noire (<i>Periparus ater</i>)	Idem Mésange bleue.
	Corvidés	Geai des chênes (<i>Garrulus glandarius</i>)	S'éloigne rarement des forêts, mais il n'est pas impossible de l'apercevoir dans un parc s'il se sent en sécurité. Il se nourrit principalement de glands.
	Sittidés	Sittelle torchepot (<i>Sitta europaea</i>)	Passereau granivore, bavard et grimpant. Elle est la seule espèce d'Europe à pouvoir descendre le long des troncs la tête en bas.
	Muscicapidés	Rouge-gorge familier (<i>Erithacus rubecula</i>)	Oiseau rondet et peu farouche. Il est identifiable grâce à son torse au plumage rouge orangé. Il est très territorial et chante toute l'année.
	Fringillidés	Pinson des arbres (<i>Fringilla coelebs</i>)	Son plumage est variable et souvent haut en couleurs pour le mâle, beaucoup plus terne pour la femelle. C'est une espèce principalement forestière qui est capable de s'adapter à de très nombreuses situations en terme d'habitat.
		Verdier d'Europe (<i>Chloris chloris</i>)	Passereau trapu aux couleurs jaunâtres. C'est une espèce très anthropophile souvent fréquente dans les parcs et jardins urbains.
		Grosbec casse-noyaux (<i>Coccothraustes coccothraustes</i>)	Espèce forestière appréciant les forêts de feuillus. Très reconnaissable par son bec imposant qui lui permet de briser la coque des graines.
	Troglodytidés	Troglodyte mignon (<i>Troglodytes troglodytes</i>)	Petit passereau au chant aigu, puissant et strident. C'est un oiseau pétulant et furtif, sans cesse en mouvement, souvent au sol ou très près du sol.

Milieu forestier	Famille	Espèce	Description
	Phylloscopidés	Pouillot véloce (<i>Phylloscopus collybita</i>)	Majoritairement insectivore, petit passereau forestier qui fréquente de nombreux autres habitats tels que les parcs et grands jardins.
	Sylviidés	Fauvette à tête noire (<i>Sylvia atricapilla</i>)	Part son dimorphisme sexuel, le mâle adulte se reconnaît à sa calotte noire alors que celle de la femelle est rousse.
	Prunellidés	Accenteur mouchet (<i>Prunella modularis</i>)	Avec son spectre d'habitat assez large, il occupe les forêts feuillues et mixtes avec une préférence pour les forêts de conifères.
	Turdidés	Grive musicienne (<i>Turdus philomelos</i>)	Espèce forestière devant disposer sur son territoire d'une zone d'alimentation au sol dégagée ou couverte de végétation herbacée
		Merle noir (<i>Turdus merula</i>)	À l'origine, espèce forestière qui a su se développer dans un grand nombre de milieux. Son chant mélodieux est audible avant le lever du soleil et au crépuscule.
Milieu aquatique et plans d'eau	Certhiidés	Grimpereau des jardins (<i>Certhia brachydactyla</i>)	Espèce arboricole qui passe la majorité de son temps accrochée à l'écorce des troncs où il trouve invertébrés et araignées pour se nourrir.

6.4 Oiseaux du milieu aquatique et des plans d'eau

TABLEAU 6 – Liste des espèces du milieu aquatique et des plans d'eau.

Milieu aquatique et plans d'eau	Famille	Espèce	Description
	Ardéidés	Héron cendré (<i>Ardea cinerea</i>)	Un des plus grands de la famille des ardéidés. Son régime se compose de poissons, d'amphibiens, et même de petits mammifères.
	Anatidés	Canard colvert (<i>Anas platyrhynchos</i>)	Canard dit "de surface" signifiant qu'il ne sait pas plonger dû à sa morphologie. Sa capacité d'adaptation est admirable et en fait le canard le plus répandu au monde.
	Rallidés	Foulque macroule (<i>Fulica atra</i>)	Se caractérise par un plumage noir avec un bec et une plaque frontale blanche. Cet oiseau est très présent dans les étangs et les points d'eau.
		Gallinule poule d'eau (<i>Gallinula chloropus</i>)	Généralement plus discret que la foulque. Sa nage est particulière, il avance penché vers l'avant en agitant nerveusement la queue.

Milieu aquatique et plans d'eau	Famille	Espèce	Description
	Motacillidés	Bergeronnette grise (<i>Motacilla alba</i>)	Reconnaissable par sa longue queue et son plumage noir et blanc. Elle occupe une large gamme d'habitats ouverts sans trop s'éloigner de points d'eau.
		Bergeronnette des ruisseaux (<i>Motacilla cinerea</i>)	Son plumage est d'un jaune intense sur les parties inférieures alors que son dos est gris. Contrairement à la B. grise, elle est très dépendante des eaux courantes où elle recherche habituellement sa nourriture
		Bergeronnette printanière (<i>Motacilla flava</i>)	Ressemble fortement à la B. des ruisseaux, s'en distingue par son plumage jaunâtre sur le dos. C'est un oiseau des milieux ouverts à semi-ouverts avec un accès facile au cours d'eau.

7 Recueil de bases de données

Ce rapport se base sur l'utilisation de plusieurs bases de données existantes. Deux sources externes de données d'oiseaux ont été exploitées. La première étant les données d'Aves récoltées lors du suivi en 2015 en format ".tif", et la seconde étant la base de données issue du site Observations.be sous format ".csv" (Annexe K). Concernant la cartographie, deux bases de données supplémentaires ont été indispensables pour la réalisation de ce mémoire : la couche SIG des écotopes venant des travaux de recherche LifeWatch sous format ".gdb", et la couche des carrés kilométriques sous format ".kml". In fine, des données personnelles ont été récoltées lors d'une campagne de terrain et enregistrées sous format ".csv" (Annexe L).

7.1 Observations.be

Observations.be (*Waarneming.be* en néerlandais) est la plus grande plateforme d'observations naturalistes en Belgique. Elle propose un module gratuit pour encoder, gérer et partager ses observations de terrain de faune et de flore avec la communauté. Ce site, né de l'association entre Natagora, Natuurpunt et des spécialistes néerlandais, sert d'outil de gestion en ligne des données naturalistes (OBSERVATION.BE, 2022). Toutes les données encodées par la communauté sont une source essentielle d'informations. En effet, le site compte désormais plus de 127 000 utilisateurs avec plus de 56 500 000 observations enregistrées depuis son lancement. Chaque année, le nombre d'observations augmentent (Figure 7).

Les principaux groupes d'espèces recensés sont les oiseaux (28 678 783), les plantes (7 790 546) et les papillons (7 747 978). Des photos et des bandes sons peuvent également être téléchargées comme preuve d'observation pour la validation automatique. À la fin du mois

de juillet 2022, 3 138 553 photos de plantes, 2 405 364 photos d’oiseaux et 2 309 942 photos de papillons de nuit sont ainsi dénombrées sur la plateforme représentant les 3 groupes les plus photographiés. Quant aux sons, moins nombreux que les photographies, 19 674 chants d’oiseaux et 665 cris de mammifères sont comptabilisés (OBSERVATION.BE, 2022).

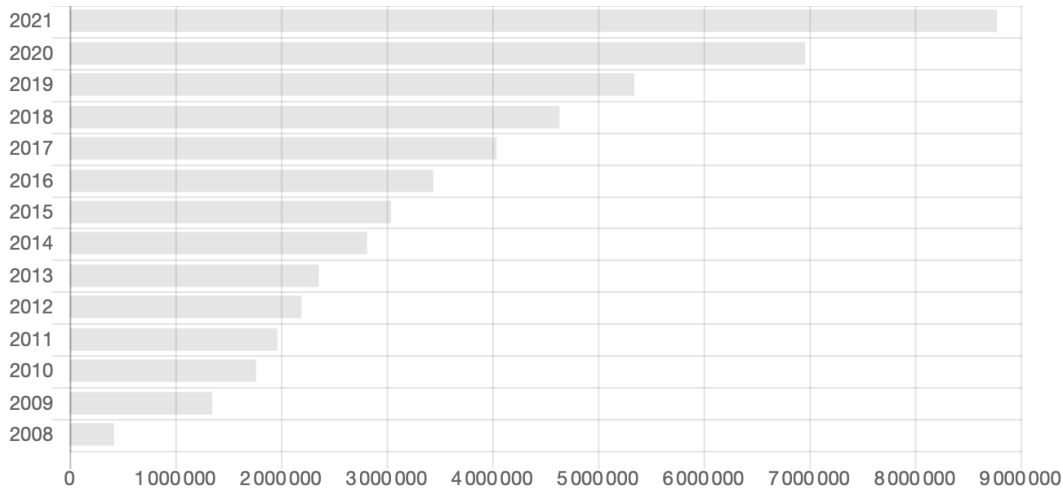


FIGURE 7 – Nombre d’observations encodées annuellement depuis 2008 sur la plateforme Observations.be, consulté en juillet 2022.

7.2 Relevés de terrain

La méthodologie retenue concernant les relevés de terrain consiste à suivre le protocole initial utilisé par Aves (cfr Section 2.5.1) lors des relevés d’oiseaux pour la conception de l’Atlas des oiseaux nicheurs. Ce qui dénote du protocole initial est que la position géographique exacte de chaque observation est précisée. Si un groupe d’individus est observé, la position géographique de l’observation correspond au centre du groupe. Comme conseillé dans l’article de SANZ-PÉREZ et al. (2020), les relevés ne sont pas réalisés si l’une des trois conditions suivantes est présente : (1) si la pluie tombe, (2) si le vent souffle à plus de 20 km/h, et (3) si les températures extérieures dépassent les 30 °C.

La prise de données a été réalisée avec deux applications mobiles. La première est *iObs* qui permet de définir l’heure de l’observation, le nombre d’oiseaux observés et de localiser spatialement chaque individu. Seules les fonctionnalités ”Choix de l’espèce”, ”Nombre”, et ”Carte” ont été utilisées. La seconde est *BirdNET*, un répertoire connecté reprenant différents chants d’oiseaux. Cette application peut ainsi analyser les fréquences des sons et d’en déterminer l’espèce émettrice.

8 Traitement des données

Plusieurs opérations ont été nécessaires avant de procéder à l'analyse des données. Tout d'abord, toutes les données localisées hors des carrés kilométriques ont été supprimées. Ensuite, les données issues du premier passage par carré kilométrique ont été nettoyées. Les espèces, présentant un effectif trop faible ou trop fluctuant dans l'échantillonnage, n'ont pas été conservées afin de rendre l'échantillonnage représentatif. C'est principalement le cas de la Rousserolle effarvate (*Acrocephalus scirpaceus*), de la Chevêche d'Athéna (*Athene noctua*) ou de la Fauvette grisette (*Sylvia communis*). Dans un deuxième temps, si plusieurs individus de la même espèce ont été observés au sein d'un même écotope, la valeur retenue est la somme de tous les individus par écotope. Concernant la précision et la validation des observations, une précision supérieure à 25 mètres n'est pas retenue. En effet, vu la petite taille possible de certains écotopes (inférieure à 10 ares), il est préférable de garder une précision suffisamment fine pour ne pas biaiser les résultats. Quant à la validation de l'observation, celle-ci est réalisée principalement par un modérateur, automatiquement ou sur base de preuves (généralement une photo ou un enregistrement audio) (Annexe M). Ces mêmes opérations sont appliquées sur les données issues du second passage par carré kilométrique et les données issues d'Observations.be pour l'année 2022.

Une fois ce nettoyage réalisé, une dernière sélection est enfin effectuée. Après les deux passages, seule la valeur de comptage la plus haute pour les individus de la même espèce au sein d'un même écotope est conservée. Une fois toutes les sources de données nettoyées, celles-ci sont rassemblées en un seul fichier.

Les manipulations cartographiques ont ensuite été réalisées avec le logiciel QGIS. Les données Excel ont été exportées en fichier ".csv" afin d'être importées et lues sur QGIS. Par une jointure de table en fonction de la localisation, les variables environnementales associées à chaque écotope ont pu être liées à chaque observation. La nouvelle table d'attributs comprenant, à présent, plusieurs informations pour chaque observation est ensuite exportée dans le logiciel R.

Avant de pouvoir utiliser une première fois le modèle, certaines variables sont dites indépendantes et doivent subir un traitement particulier. C'est principalement le cas des variables composées de proportions de sol à 250 mètres de l'écotope. En effet, elles ne peuvent être intégrées directement dans le modèle en raison de la colinéarité entre celle-ci.

Dans une régression, la colinéarité entre variables est présente lorsque ces variables sont fortement corrélées entre elles. "Les conséquences de la colinéarité statistique sont les suivantes

(FOUCART, 2006) :

- Les coefficients de régression estimés peuvent être élevés en valeur absolue ;
- Leurs signes peuvent être contraires à l'intuition ;
- Les variances des estimateurs peuvent être élevées ;
- Les coefficients de régression et le coefficient de corrélation multiple sont instables par rapport aux coefficients de corrélation entre les variables explicatives.”

Il n'est donc pas possible d'utiliser ces variables telles quelles puisque la colinéarité crée des difficultés importantes dans l'interprétation des résultats. Un « recodage » de ces variables (λ_i) doit alors être réalisé en choisissant une variable de référence π_n . C'est la variable Plowed250 qui a été sélectionnée comme variable de référence. Plowed250 représente la proportion d'occupation de terres arables (cultures annuelles et couvert herbacé temporaire) à 250 mètres de rayon de l'écotope sélectionné. Cette variable a été définie comme variable de référence car elle est la plus répandue sur le territoire d'étude, représentant ainsi la part la plus importante de la matrice d'occupation de sol.

Soit une variable composée de proportion telle que :

$$\pi_i \in [0, 1] \forall i = 1, \dots, n \quad \text{avec} \quad \sum_i^n \pi_i = 1 \quad (1)$$

La variable (λ_i) est définie de la manière suivante :

$$\lambda_i = \log\left(\frac{\pi_i}{\pi_n}\right) = \log(\pi_i) - \log(\pi_n) \quad \forall i = 1, \dots, n \quad (2)$$

Et donc,

$$\lambda_n = \log\left(\frac{\pi_n}{\pi_n}\right) = \log(1) = 0 \quad (3)$$

Les variables (λ_1), ..., (λ_n) sont donc intégrées dans la base d'attributs selon une échelle logarithmique et peuvent ainsi être lues par le modèle sans qu'il y ait des problèmes de colinéarité entre elles.

Une opération mathématique est encore à effectuer pour pallier un dernier problème. En effet, il se peut qu'une des proportions de sol à 250 mètres de l'écotope soit nulle. Le facteur π_i équivaut dans ce cas à zéro. Si on l'intègre tel quel dans l'équation 2 ci-dessus, le logarithme de $(\pi_i)/(\pi_n)$ retournera une erreur puisque $\log(0)$ n'est pas défini. Étant donné que les zéros ne peuvent être considérés comme tels, ils peuvent dès lors être considérés comme des « faux » zéros résultant d'arrondis. Si les proportions sont mesurées au pourcent près, une solution

est alors de remplacer chaque 0 par 0,0001⁷ et de décaler des autres catégories les ajouts dans des proportions identiques pour que le total fasse toujours 100 %.

Un exemple est illustré avec le cas suivant où la base de données comprend 5 variables explicatives d'une proportion de surface :

- Variable 1 : 0
- Variable 2 : 0
- Variable 3 : 0,2
- Variable 4 : 0,5
- Variable 5 : 0,3

La somme des 5 variables vaut 1 signifiant que 100 % de la surface est répartie en plusieurs catégories d'occupation du sol par exemple. Cependant, il y a la présence de zéros qui ne peuvent pas être introduits dans l'équation logarithmique. Les variables sont modifiées comme telles :

- Variable 1 : $0 + 0,0001 = 0,0001$
- Variable 2 : $0 + 0,0001 = 0,0001$
- Variable 3 : $0,2 - (0,0001*2)/3 = 0,19993$
- Variable 4 : $0,5 - (0,0001*2)/3 = 0,49993$
- Variable 5 : $0,3 - (0,0001*2)/3 = 0,29993$

De cette manière, les zéros ont été éliminés de la base de données et la somme des variables est toujours égale à 1 sans que cela ait trop d'impact sur les résultats. Les nouvelles valeurs peuvent dès lors être introduite dans les équations logarithmiques en prenant la variable 3 comme variable de référence par exemple. Les nouvelles variables ainsi créées suivent une échelle logarithmique comme montré ci-dessous :

- logVariable 1 : $\log (0,0001/0,19993) = -3,3008$
- logVariable 2 : $\log (0,0001/0,19993) = -3,3008$
- logVariable 3 : $\log (0,19993/0,19993) = 0$
- logVariable 4 : $\log (0,49993/0,19993) = 0,3981$
- logVariable 5 : $\log (0,29993/0,19993) = 0,1761$

Finalement, ce sont ces nouvelles valeurs qui sont attribuées aux variables de proportion de sol à 250 mètres de rayons de l'écotopie. Celles-ci permettent alors de pallier le problème de colinéarité entre ces variables explicatives.

Le Tableau 7 reprend l'ensemble des attributs finaux pour chaque espèce.

7. Cette valeur a été choisie de façon à correspondre au plus à zéro pour que la variable ne prenne pas trop d'importance dans le modèle.

TABLEAU 7 – Liste des variables de chaque espèce.

Variable	Description
Nombre	Nombre d'individus observés
x	Coordonnée X lambert de l'observation
y	Coordonnée Y lambert de l'observation
Artif	Surface de zone bâtie [m ²]
BroadLV	Surface de forêt feuillue [m ²]
NeedlLV	Surface de forêt de résineux [m ²]
Plowed	Surface de terres arables (cultures annuelles et couvert herbacé temporaire) [m ²]
MGramin	Surface de couverture permanente de graminoides monospécifiques [m ²]
NOpen	Surface de zone ouverte avec des sols relativement secs [m ²]
Bare	Surface de sols nus ou peu couverts de végétation (<15%) [m ²]
Water	Surface de plans d'eau permanents [m ²]
COpen	Surface de couvert herbacé et arbustif permanent de plantes adventices forestières et de jeunes arbres [m ²]
Autre	Surface des résidus de pourcentage [m ²]
elev_mean	Altitude moyenne de l'écotop [m]
Sun_Spring	Énergie lumineuse moyenne annuelle [W/m ²]
Dst_Road	Distance à la route de l'écotop [m]
Dst_Rail	Distance à la voirie ferrée de l'écotop [m]
Dst_River	Distance à la rivière de l'écotop [m]
Dst-Sea	Distance à la mer de l'écotop [m]
Dst_Forest	Distance à la forêt de l'écotop [m]
Shape_Area	Surface totale de l'écotop [m ²]
logArtif250	Valeur corrigée pour la proportion de zone bâtie à l'écotop sur un rayon de 250 mètres
logBraodLV250	Valeur corrigée pour la proportion de forêt feuillue à l'écotop sur un rayon de 250 mètres
logNeedlLV250	Valeur corrigée pour la proportion de forêt de résineux à l'écotop sur un rayon de 250 mètres
logPlowed250	Valeur corrigée pour la proportion de terre arable à l'écotop sur un rayon de 250 mètres
logMGramin250	Valeur corrigée pour la proportion de couverture monospécifique à l'écotop sur un rayon de 250 mètres
logNOpen250	Valeur corrigée pour la proportion de zone ouverte avec sol sec à l'écotop sur un rayon de 250 mètres
logBare250	Valeur corrigée pour la proportion de sol nu ou peu couvert à l'écotop sur un rayon de 250 mètres
logWater250	Valeur corrigée pour la proportion de plan d'eau à l'écotop sur un rayon de 250 mètres
logCOpen250	Valeur corrigée pour la proportion de couverture herbacée à l'écotop sur un rayon de 250 mètres
logWOpen250	Valeur corrigée pour la proportion de zone ouverte avec sol humide à l'écotop sur un rayon de 250 mètres
logAutre250	Valeur corrigée pour la proportion du pourcentage résiduel à l'écotop sur un rayon de 250 mètres

8.1 Méthodes d'analyse pour la prédiction des densités des populations

La densité des populations d'oiseaux sera prédite, dans ce travail, par une régression de Poisson. Cette régression convient parfaitement lorsque la variable cible Y est une variable de comptage (CONSUL & FAMOYE, 1992). La fonction de probabilité de la loi de Poisson $P(Y = y)$ est donnée par (HILBE, 2011) :

$$P(Y = y) = \frac{e^{-\lambda} \lambda^y}{y!} \quad (4)$$

Où,

- $P(Y = y)$ est la probabilité de l'occurrence de y ;
- y est le nombre d'évènements observés ;
- λ est le paramètre de la loi.

La particularité du modèle réside dans son paramètre λ . En effet, λ est l'espérance et la variance. Il peut donc être considéré comme le nombre attendu de y . L'objectif de la régression de Poisson est donc de modéliser ce paramètre λ à l'aide de covariables $(X_1, X_2, \dots, X_j)$ via une régression linéaire simple, avec ("STAT 501 - Regression Methods", s. d.) :

$$\ln(\lambda) = a_0 + a_1 X_1 + \dots + a_j X_j \quad (5)$$

Ces covariables sont principalement les variables environnementales définies par écotope (occupation du sol, topographie, type de sols, temps d'ensoleillement, etc.). Pour chaque espèce donnée, la meilleure combinaison de ces covariables a été sélectionnée sur base du critère d'information d'Akaike (AIC) ((CAVANAUGH & NEATH, 2019)). Une vérification automatique, par un pseudo R-carré est utilisée pour l'auto-corrélation des variables et pour l'ajustement du modèle. Une fois le modèle correctement calibré, l'abondance de cette espèce observée par écotope peut être calculée. L'estimation de l'abondance une fois évaluée, les résultats sont liés à la couche SIG des écotopes par leur identifiant unique. Ensuite, en appliquant une symbologie graduée, la cartographie de l'abondance de l'espèce peut être réalisée.

9 Validation des résultats

Une fois l'abondance des espèces estimée par écotope, chaque cartographie a été comparée aux cartes obtenues par Aves en 2015 afin de juger si les zones de fortes et basses densités coïncident entre elles. Si c'est le cas, la cartographie obtenue par notre modèle est estimée représentative de la réalité de terrain. Dans le cas contraire, elle est estimée non

représentative de la réalité de terrain. Une fois la comparaison réalisée, une évaluation de la qualité du modèle est requise. Plusieurs manipulations QGIS (Figure 8) vont se succéder afin de pouvoir calculer un R^2 . Celui-ci est évalué en faisant le rapport des estimations prédites en 2015 par les abondances estimées du modèle testé dans cette étude.

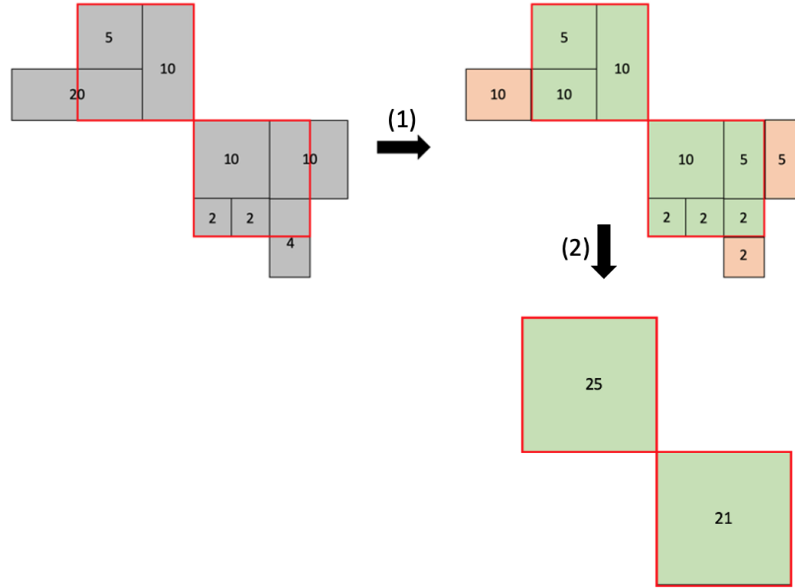


FIGURE 8 – Succession des différentes étapes d’agrégation des abondances à 1 kilomètre de résolution. 1) Calcul de l’abondance par polygone (délimités en noir) proportionnellement à la surface de chaque polygone ; 2) Calcul de la somme des abondances au sein des carrés kilométriques (délimités en rouge).

Dans un premier temps, il faut agréger les résultats obtenus à l’échelle des carrés kilométriques. Pour cela, une intersection entre la couche ”écotope” et la couche ”carré kilométrique” est effectuée. De nouveaux polygones sont alors créés. Un champ surface est ajouté dans la table attributaire des nouveaux polygones. L’abondance estimée du polygone initial est calibrée de manière proportionnelle à la surface des nouveaux polygones créés selon l’équation (6).

$$A_f = \frac{S_f}{S_i} * A_i \quad (6)$$

Où,

- A_f est l’abondance finale calculée par nouveau polygone
- S_f est la surface finale du nouveau polygone
- S_i est la surface initiale du polygone avant sa division
- A_i est l’abondance initiale calculée par écotope

Dans un deuxième temps, la somme des abondances calculées de chaque polygone constituant un carré kilométrique est comptée. Ainsi, l'abondance de chaque espèce est agrégée à un kilomètre de résolution.

Dans un troisième et dernier temps, le rapport entre ces abondances agrégées et l'abondance kilométrique pour l'année 2015, estimée par Aves, permet de mesurer un R^2 . Celui-ci évalue la qualité des résultats obtenus. Pour mesurer la validation des modèles, 20 carrés kilométriques, répartis de manière aléatoire au sein de la zone d'étude, ont été utilisés.

10 Analyse de sensibilité

L'analyse de sensibilité permet d'évaluer l'impact des modifications de certaines variables ou méthode de calcul sur le résultat. Plusieurs variantes sont testées et un nouveau R^2 est calculé afin d'évaluer si la modification a un apport positif ou négatif sur le résultat initial. Cinq variantes ont été étudiées et sont présentées ci-dessous.

10.1 Utilisation des données issues de la campagne de terrain uniquement sans les données provenant du site Observations.be

L'utilisation des données issues des sciences participatives permet d'augmenter le nombre de données, contrainte principale de cette étude. Cependant, le recensement des espèces par des observateurs non professionnels ne suit pas un protocole officiel, ce qui aboutit à des données mal enregistrées. Dès lors, il se peut qu'une partie des données encodées soient des données dites "de présence". Cela s'explique par le fait que les observateurs n'informent pas nécessairement le nombre d'individus observés. À titre d'exemple, un groupe de mésanges est aperçu par un observateur qui enregistre le nombre "1" de mésanges dans son application ou sur le site, informant alors uniquement la présence de l'espèce et non le nombre d'individus repérés. L'utilisation conjointe des données de présence et des données de comptage dans le modèle pourrait mener un biais dans les résultats. Pour cette raison, il est jugé intéressant de voir si le modèle est amélioré lorsque seules les données issues de la campagne de terrain sont utilisées.

10.2 Ajout des données d'absence

Les données issues des sciences participatives ne suivent pas un protocole officiel et informent les données de présence des espèces comme expliqué à l'instant. Dès lors, celles-ci sont également utiles pour informer les données "d'absence". En effet, le protocole officiel

force l'observateur à suivre un chemin prédéfini pendant une durée déterminée lors de la sortie de terrain. Grâce au recensement de ces observateurs, de nombreuses zones du territoire peuvent également être prospectées. Par conséquent, si une espèce n'a pas été observée dans une région du territoire, ni par un observateur suivant le protocole officiel, ni par un observateur amateur, l'espèce est alors considérée absente dans ce secteur et dans cet écotope. L'ajout de ces données d'absence permettrait de calibrer au mieux les prédictions du modèle.

10.3 Utilisation d'une variable de référence différente lors du problème de colinéarité entre variables

Lors du traitement des données (cfr Section 8), certaines variables indépendantes doivent être traitées pour pallier le problème de colinéarité entre variables. C'est principalement le cas pour les variables composées de proportions de sol à 250 mètres de rayon de l'écotope. Pour permettre de traiter ces données, une variable de référence est choisie, la variable *Plowed250*. Cette dernière a été sélectionnée car elle est la plus répandue dans la zone d'étude et représente ainsi correctement la matrice d'occupation du sol. Cependant, une autre variable aurait pu être utilisée, en altérant ou pas les résultats.

10.4 Utilisation d'un autre modèle de type "Quasi-Poisson"

Les données de comptage sont généralement sujettes à une dispersion, ce qui induit une variabilité plus importante dans un modèle statistique (MACKENZIE J. EDMONDSON et al., 2022). Pour pallier cette difficulté, un modèle de type "Quasi-Poisson" peut être adopté. Cependant, la distribution Quasi-Poisson est sujette à des complications concernant sa modélisation et son choix de variables avec le programme R. En effet, ce logiciel ne permet pas de sortir ni une estimation de vraisemblance, ni une valeur AIC pour ce type de distribution. Dès lors, l'ensemble des variables a été déterminé de façon manuelle par une suppression itérative des variables en fonction de la pertinence de celles-ci et de leur p-valeur respective.

10.5 Caractérisation d'une zone de territoire par groupes d'individus lors du calcul du R^2

Une méthode a été élaborée (cfr Section 9) afin de valider les résultats grâce au calcul d'un R^2 . Cette méthode pourrait être modifiée en distinguant des zones de territoire par groupe d'individus. En effet, les oiseaux sont des animaux très mobiles qui occupent souvent des territoires plus vastes que certains écotopes. Lorsque le modèle prédit le même nombre d'individus (ou une même gamme d'abondance) dans plusieurs écotopes conjoints disposant de

variables explicatives fortement similaires, il se pourrait que ces écotopes représentent le territoire d'un seul groupe d'individus et non d'un groupe d'individus par écotope. Une nouvelle méthode de validation consiste alors à calculer l'abondance par polygone proportionnellement à la surface de chaque polygone comme expliqué à la Section 9 et d'ensuite joindre les polygones conjoints ayant la même gamme d'abondance. Ces abondances sont sommées au sein des carrés kilométriques pour terminer. La succession des étapes est présentée à la Figure 9.

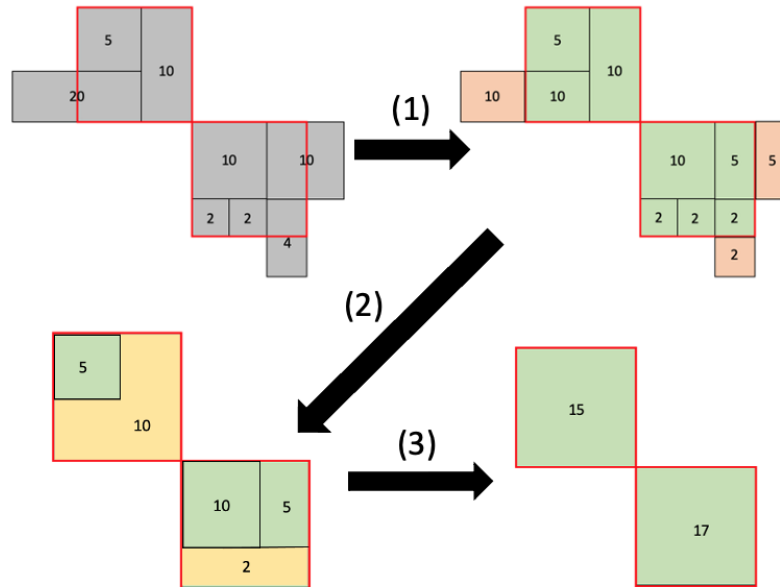


FIGURE 9 – Succession des différentes étapes d'agrégation des abondances à 1 kilomètre de résolution. 1) Calcul de l'abondance par polygone (délimités en noir) proportionnellement à la surface de chaque polygone; 2) Détermination des zones de territoire par la fusion de polygones conjoints ayant la même gamme d'abondance; 3) Calcul de la somme des abondances au sein des carrés kilométriques (délimités en rouge).

Résultats

11 Caractéristiques générales des données

Durant la campagne de terrain de cette étude, du 3 mars 2022 au 17 mai 2022, 2289 individus ont été recensés. On y compte 41 espèces distinctes réparties dans les quatre milieux décrits plus haut. Le mois de mars comptabilise le plus d'observations (921), suivi par le mois d'avril (757) puis par le mois de mai (611).

De manière générale, la précision et la validation des données est bonne. En effet, la précision moyenne s'élève à 10,7 mètres et seulement 3,6 % des données sont considérées incertaines après une validation automatique (statut de validation disponible à l'Annexe M). Après un nettoyage des données (en éliminant les précisions supérieures à 25 mètres, les données rejetées par la validation et les données localisées hors des quatorze carrés kilométriques), c'est finalement 2099 individus qui sont retenus pour l'étude.

Concernant les données issues de la plateforme *Observations.be*, 28 251 observations, dont 154 espèces différentes, ont été enregistrées au sein de la zone d'étude depuis le 1^{er} janvier 2022 par 141 observateurs différents. Le taux d'observation est plus élevé pour le mois de janvier. La Figure 10 présente la distribution des observations enregistrées sur la plateforme ainsi que sur le terrain au cours des cinq premiers mois de l'année 2022. La donnée de la plateforme la plus récente date du 24 avril sur le site de Pécrot (petit village de la commune de Grez-Doiceau). Le mois de janvier comptabilise le plus grand nombre d'observations (12 854).

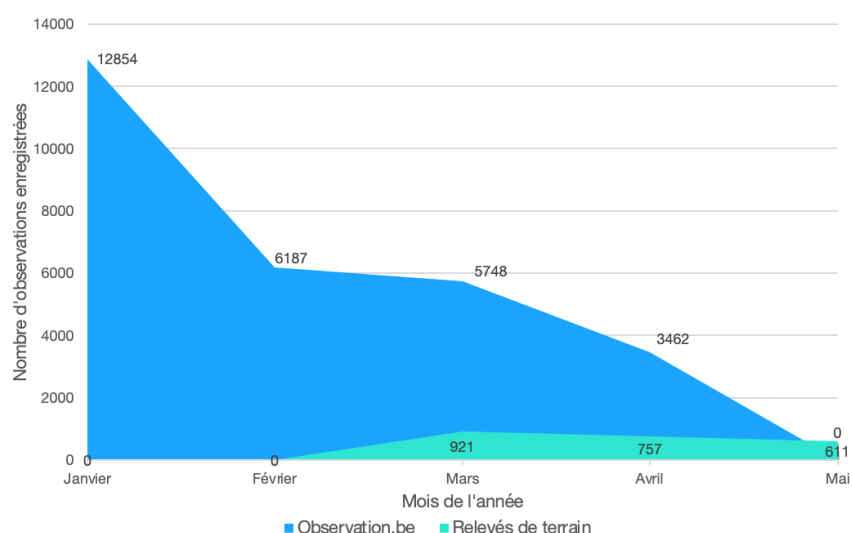


FIGURE 10 – Distribution des observations enregistrées sur la plateforme *Observations.be*, ainsi que les relevés de terrain au cours des cinq premiers mois de l'année 2022.

La précision des données brutes d'Observation.be varie entre 1 et 999 mètres. Une précision de 999 mètres est généralement associée au fait que l'observateur a enregistré ses données une fois rentré chez lui et non sur le terrain. Cependant, cela n'est appliqué que pour 4 observations. En excluant ces données, la précision maximale est alors de 120 mètres. La précision moyenne de l'ensemble des données issues de la plateforme est de 13,2 mètres. Concernant la validation des données, 82 % ont été acceptées (A), 2% ont été acceptées avec une preuve (J), 16 % sont encore indéterminées et une seule a été rejetée (N). Une fois les données également nettoyées, 2903 individus (82 espèces différentes) sont comptabilisées.

Parmi ces 82 espèces, seules les 41 espèces observées lors des relevés de terrain sont ajoutées à la base de données finale, soit 1934 individus. Au total, en sommant les données issues des 41 espèces des relevés de terrain et celles provenant du site *Observations.be*, ce sont 4033 observations d'oiseaux, réalisées dans les 14 carrés kilométriques de la zone d'étude et réparties dans les quatre milieux, qui sont utilisées pour le modèle. Pour chacune de ces espèces, le modèle a tenté de prédire leur abondance par écotope. Les résultats sont donc présentés sous forme de cartographie de la densité par écotope de chaque espèce en question.

Pour rappel (cfr Section 9), les cartographies obtenues sont comparées avec celles réalisées par Aves afin d'estimer si elles sont représentatives ou non de la réalité de terrain, et ce, avant le calcul de leur validation.

12 Cartographies représentatives de la réalité de terrain

La Mésange charbonnière (*Parus major*) et le Pinson des arbres (*Fringilla coelebs*) sont les deux espèces pour lesquelles les cartographies sont estimées représentatives de la réalité de terrain. Le Tableau 8 reprend le nombre d'individus enregistrés dans la base de données ainsi que la validation de chacun des résultats.

TABLEAU 8 – Nombre d'observations des espèces dont le résultat est représentatif de la réalité de terrain et leur validation cartographique.

Nom de l'espèce	Nombre d'individus	Validation (R^2)
Mésange charbonnière (<i>Parus major</i>)	371	0,28
Pinson des arbres (<i>Fringilla coelebs</i>)	235	0,25

Ayant le R^2 le plus élevé de cette catégorie, la répartition de l'abondance de la M. charbonnière est illustrée à la Figure 11. La cartographie obtenue pour le P. des arbres est quant à elle visible à l'Annexe N.

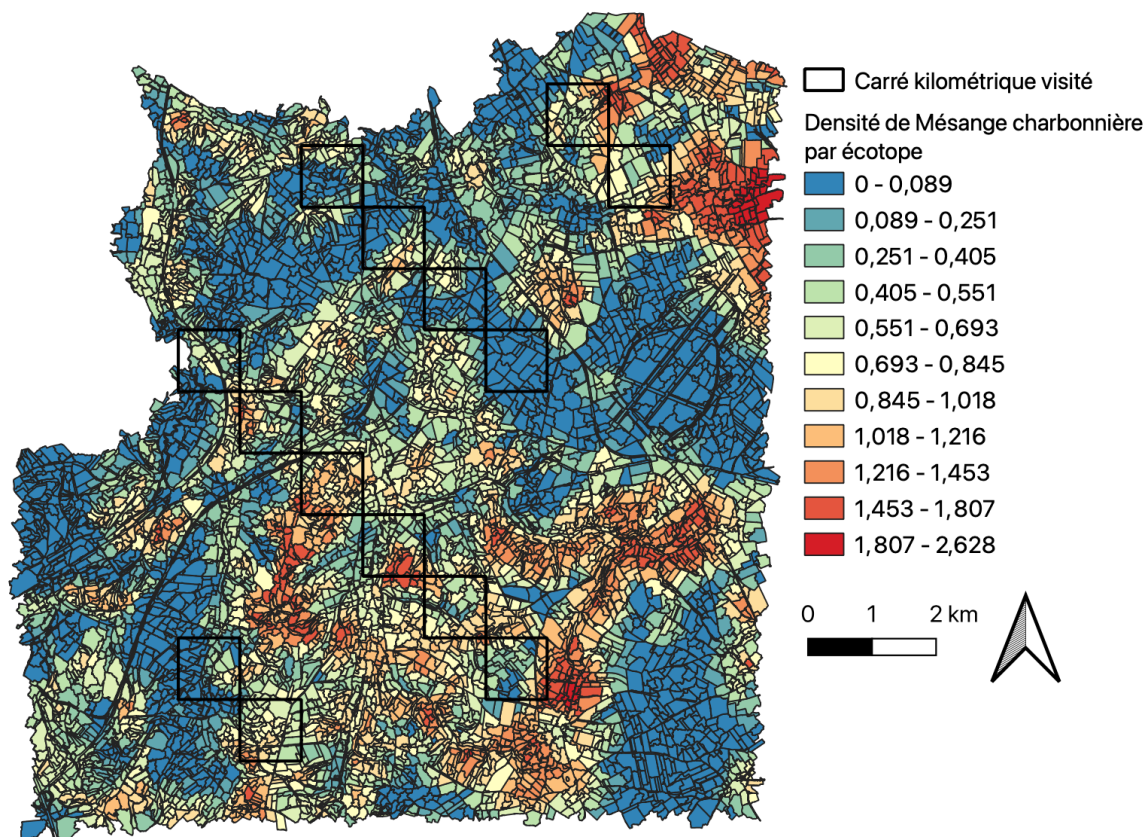


FIGURE 11 – Répartition de l'abondance de la Mésange charbonnière par écotope.

13 Cartographies non représentatives de la réalité de terrain

Les cartographies de sept autres espèces ont été estimées non représentatives de la réalité de terrain. Le Tableau 9 reprend le nombre d'individus enregistrés dans la base de données ainsi que la validation de chacun des résultats.

TABLEAU 9 – Nombre d'observations des espèces dont le résultat est non représentatif de la réalité de terrain et leur validation cartographique.

Nom de l'espèce	Nombre d'individus	Validation (R^2)
Corneille noire (<i>Corvus corone</i>)	442	0,15
Choucas des tours (<i>Coloeus monedula</i>)	430	0,04
Moineau domestique (<i>Passer domesticus</i>)	340	0,07
Pigeon ramier (<i>Columba palumbus</i>)	330	0,07
Alouette des champs (<i>Alauda arvensis</i>)	164	0,23
Tourterelle turque (<i>Streptopelia decaocto</i>)	76	0,08
Fauvette à tête noire (<i>Sylvia atricapilla</i>)	73	0,22

Ayant le R^2 le plus élevé de cette catégorie, la répartition de l'abondance de l'Alouette des champs est illustrée à la Figure 12. Les cartographies obtenues pour l'ensemble des six autres espèces sont visibles à l'Annexe O.

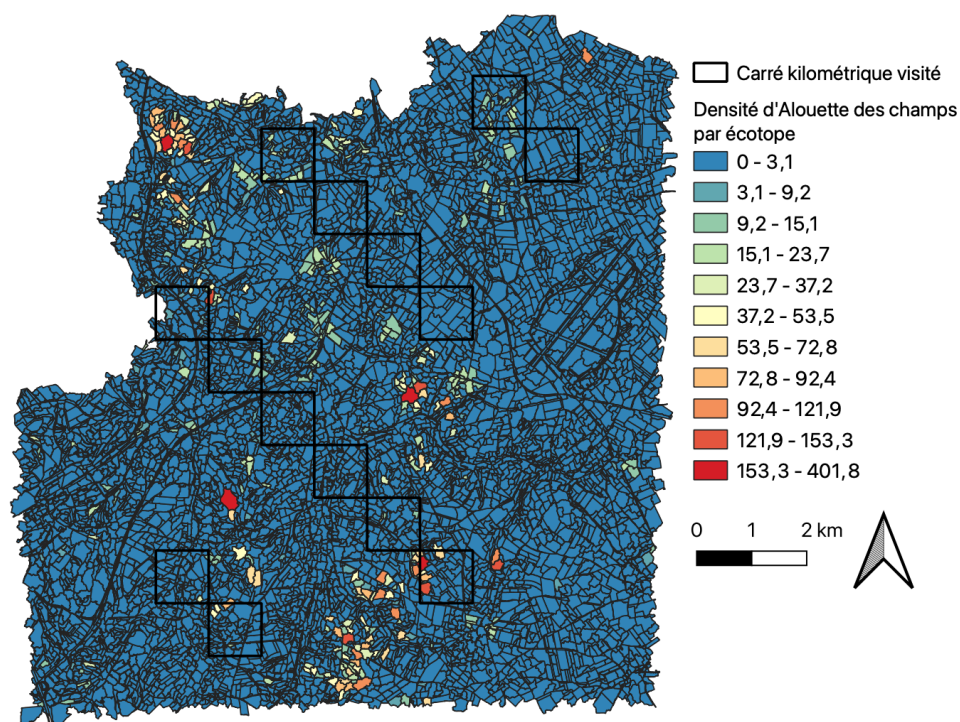


FIGURE 12 – Répartition de l'abondance de l'Alouette des champs par écotope.

14 Détermination de l'intercept

Pour les 32 espèces qui n'ont pas encore été présentées dans les résultats, le modèle n'a pas pu définir un jeu de variables explicatives permettant de cartographier l'abondance de ces espèces par écotopes. Le modèle a simplement identifié l'intercept pour ces espèces. Dès lors, la comparaison ainsi que la validation des cartographies n'a pas été réalisée. Le tableau suivant reprend le nombre d'individus enregistrés dans la base de données.

TABLEAU 10 – Nombre d'observations des espèces dont le modèle identifie l'intercepte. Le "-" signifie que la validation n'a pas été réalisée car aucune cartographie n'a été obtenue.

Nom de l'espèce	Nombre d'individus	Validation (R^2)
Mésange bleue (<i>Cyanistes caeruleus</i>)	192	-
Troglodyte mignon (<i>Troglodyte troglodyte</i>)	173	-
Rouge-gorge familier (<i>Erithacus rubecula</i>)	161	-
Pie bavarde (<i>Pica pica</i>)	125	-
Étourneau sansonnet (<i>Sturnus vulgaris</i>)	121	-
Pouillot véloce (<i>Phylloscopus collybita</i>)	109	-
Merle noir (<i>Turdus merula</i>)	94	-
Accenteur mouchet (<i>Prunella modularis</i>)	67	-
Canard colvert (<i>Anas platyrhynchos</i>)	67	-
Geai des chênes (<i>Garrulus glandarius</i>)	58	-
Bruant jaune (<i>Emberiza citrinella</i>)	53	-
Verdier d'Europe (<i>Chloris chloris</i>)	40	-
Pic épeiche (<i>Dendrocopos major</i>)	37	-
Bergeronnette grise (<i>Motacilla alba</i>)	32	-
Sittelle torchepot (<i>Sitta europaea</i>)	29	-
Grive musicienne (<i>Turdus philomelos</i>)	26	-
Pigeon biset domestique (<i>Columba livia domestica</i>)	25	-
Pic vert (<i>Picus viridis</i>)	23	-
Buse variable (<i>Buteo buteo</i>)	20	-
Rougequeue noir (<i>Phoenicurus ochruros</i>)	18	-
Grimpereau des jardins (<i>Certhia brachydactyla</i>)	16	-
Gallinule poule d'eau (<i>Gallinula chloropus</i>)	15	-
Linotte mélodieuse (<i>Linaria cannabina</i>)	15	-
Bergeronnette printanière (<i>Motacilla flava</i>)	10	-
Faucon Crécerelle (<i>Falco tinnunculus</i>)	10	-
Héron cendré (<i>Ardea cinerea</i>)	9	-
Foulque macroule (<i>Fulica atra</i>)	9	-
Faisan de Colchide (<i>Phasianus colchicus</i>)	9	-
Grosbec casse-noyaux (<i>Coccothraustes coccothraustes</i>)	7	-
Mésange noire (<i>Periparus ater</i>)	6	-
Bergeronnette des ruisseaux (<i>Motacilla cinerea</i>)	4	-
Épervier d'Europe (<i>Accipiter nisus</i>)	0	-

15 Catégorisation des résultats sous forme d'une clé dichotomique

L'analyse de l'ensemble des résultats a permis d'identifier une clé dichotomique permettant de catégoriser les différents résultats obtenus. Cette clé, ainsi que les valeurs de sélection, ont été déterminées une fois tous les résultats obtenus (Figure 13).

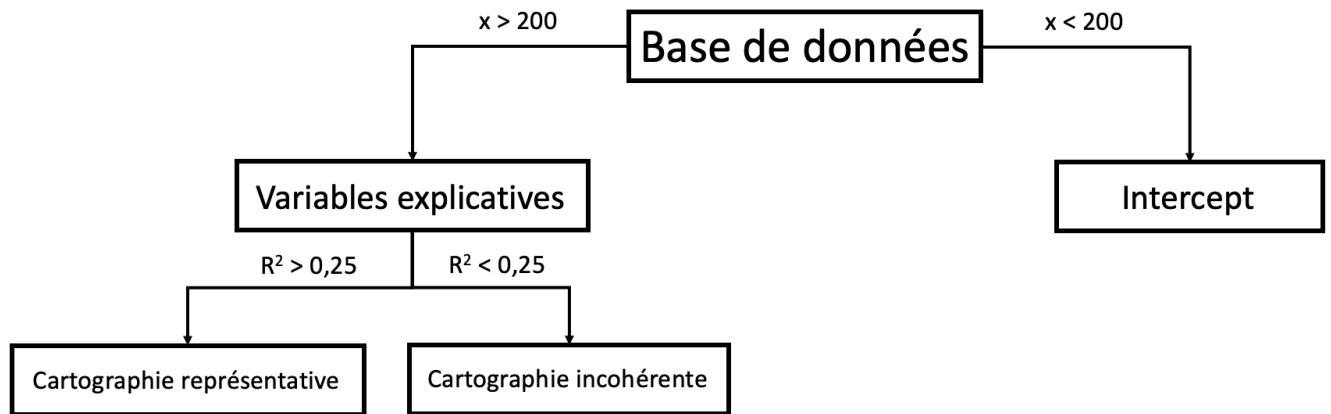


FIGURE 13 – Clé dichotomique catégorisant les différents résultats obtenus, où x est le nombre d'individus enregistrés et R^2 est le résultat de la validation.

Le premier niveau de sélection permet de déterminer si le modèle est capable de définir un jeu de variables explicatives en fonction du nombre d'individus enregistrés dans la base de données. Un seuil minimum de 200 individus observés est nécessaire au modèle pour que celui-ci définisse un jeu de variables explicatives et puisse cartographier l'abondance par écotope de ces individus. En dessous de ce seuil, le modèle identifie uniquement l'intercept. Ceci est observable pour l'ensemble des espèces à l'exception de trois d'entre elles : l'Alouette des champs, la Tourterelle turque et la Fauvette à tête noire, avec respectivement 164, 76 et 73 individus dénombrés. Cependant, malgré le fait que le modèle a défini un jeu de variables, les cartographies de ces trois espèces ne sont pas estimées représentatives de la réalité de terrain.

Le second niveau de sélection permet d'estimer si la cartographie est représentative ou non de la réalité de terrain, sur base de la valeur de la validation. Dans le cadre de cette étude, il a été jugé qu'au-dessus d'un R^2 de 0,25, la cartographie est considérée comme représentative de la réalité de terrain. Cette affirmation est soutenue par la comparaison des trois tableaux reprenant les valeurs de R^2 des différentes espèces (les cartographies représentatives de la réalité de terrain sont supérieures ou égales à 0,25, et inversement).

Pour conclure, les résultats peuvent être divisés en trois catégories suivant cette clé dichotomique :

1. Les espèces pour lesquelles le modèle identifie un jeu de variables explicatives cohérentes et produit une cartographie représentative de la répartition des populations de l'espèce en question (lorsque le R^2 est supérieur à 0,25) ;
2. Les espèces pour lesquelles le modèle identifie un jeu de variables explicatives incohérentes et produit une cartographie non représentative de la répartition des populations de l'espèce en question (lorsque le R^2 est inférieur à 0,25) ;
3. Les espèces pour lesquelles le modèle ne peut pas identifier un jeu de variables explicatives en raison d'un manque de données (x inférieur à 200 individus enregistrés).

16 Tests de sensibilité

16.1 Utilisation des données de terrain uniquement

Le test d'utilisation des données de terrain a été effectué sur la Mésange charbonnière, car elle compte le plus de données issues de la campagne de terrain sans prendre en compte les données issues du site Observations.be. Cette espèce comptabilise alors 123 individus enregistrés lors des sorties.

Le modèle n'a pas pu identifier un jeu de variables explicatives, ce qui est en adéquation avec la clé dichotomique puisque le nombre d'individus enregistrés est inférieur au seuil minimal des 200 observations nécessaires. Dès lors, le modèle ne peut déterminer que l'intercept.

16.2 Ajout des données d'absence

Le nombre d'individus enregistrés d'Alouette des champs est inférieur au seuil minimal requis (164). Faisant exception à la règle, le modèle est tout de même parvenu à prédire des abondances par écotopes, comme observé dans les résultats. Cependant, ces prédictions ne sont pas bien calibrées et ne permettent pas de cartographier correctement les densités de l'espèce sur la zone d'étude (cfr Section 13).

Cependant, lorsque des données d'absences sont ajoutées dans la base de données de l'Alouette des champs, cela abouti à une meilleure calibration du modèle. Ce dernier identifie alors un jeu de variables explicatives permettant de mieux estimer l'abondance d'Alouette par écotope. En effet, la validation est passée d'un R^2 de 0,23 à 0,32 (Figure 14).

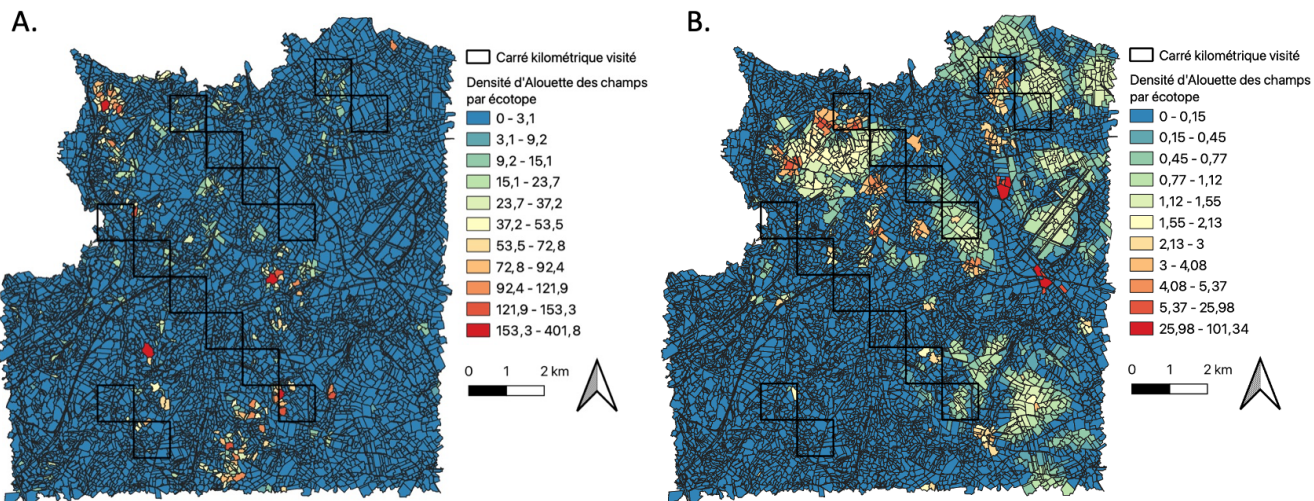


FIGURE 14 – Amélioration de la calibration du modèle par l’ajout de données d’absence. A. Cartographie sans les données d’absence ($R^2 = 0,23$) ; B. Cartographie avec les données d’absence ($R^2 = 0,32$).

Quelques anomalies restent présentes sur la nouvelle carte. Par exemple, l’estimation de la densité d’Alouette est fortement supérieure à la moyenne pour cinq écotopes. Après vérification, la raison de ces erreurs s’explique par la grande proportion de surface de la variable ”Autre” dans ces cinq écotopes, qui affecte la prédiction du modèle.

16.3 Utilisation d’une variable de référence différente

La variable de référence choisie pour chaque espèce est la variable Plowed250. À titre d’exemple, trois régressions utilisant des variables de référence différentes (BroadLV250, WOpen250 et COpen250) ont été testées pour modéliser l’abondance du Pinson des arbres (Figure 15). Cette figure montre que le choix de la variable de référence a un impact important sur les résultats. En effet, selon le choix de cette variable de référence, les variables explicatives sélectionnées par le modèle ne sont pas les mêmes, modifiant les zones de haute et basse densité de l’espèce et les gammes d’abondances, ces dernières variant de 0,98 à 1,93 pour les densités maximales par écotope.

Les valeurs de validation de chacune de ces cartes montrent qu’un meilleur résultat est obtenu en utilisant la variable Plowed250 comme variable de référence. En effet, la carte A obtient un R^2 de 0,25, ce qui la catégorise comme représentative de la réalité de terrain (catégorie 1 suivant la clé dichotomique). A contrario, les trois autres cartographies sont catégorisées comme non représentative (catégorie 2 suivant la clé dichotomique) puisqu’elles obtiennent un R^2 inférieur à 0,25.

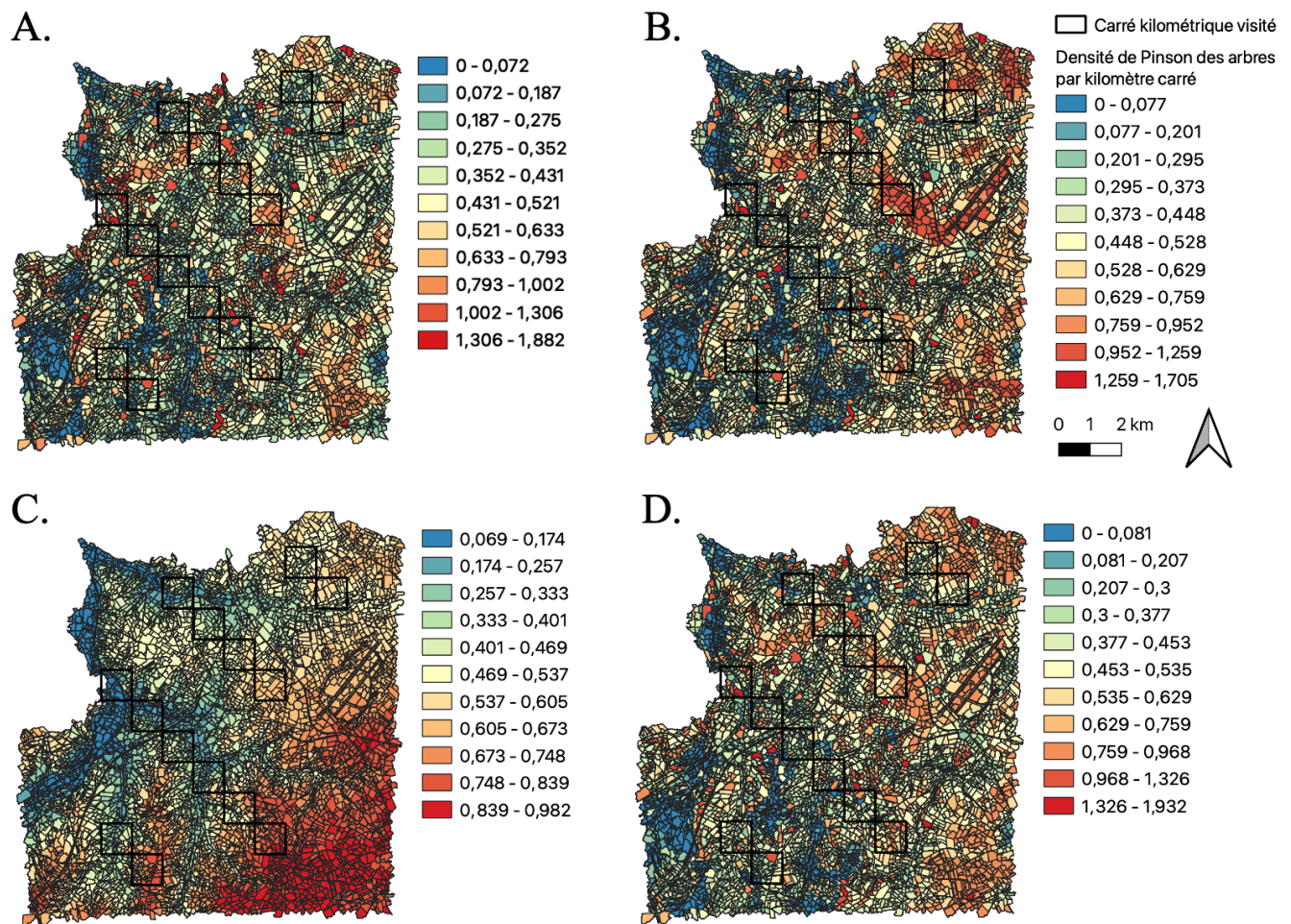


FIGURE 15 – Influence du choix de la variable de référence sur les résultats (cas du Pinson des arbres). A. Plowed250 est choisi comme variable de référence ($R^2 = 0,25$) ; B. BroadLV250 est choisi comme variable de référence ($R^2 = 0,18$) ; C. WOpen250 est choisi comme variable de référence ($R^2 = 0,15$) ; D. COpen250 est choisi comme variable de référence ($R^2 = 0,18$).

16.4 Modélisation avec une régression Quasi-Poisson

La régression Quasi-Poisson a été expérimentée sur la Mésange bleue. Cette espèce compte 192 individus enregistrés. Le modèle n'a alors identifié aucune variable explicative puisque le seuil des 200 observations n'a pas été atteint.

Or, l'utilisation d'une régression de type Quasi-Poisson a permis de déterminer un jeu de variables et le paramètre de dispersion (0,4). Grâce à cela, l'abondance de la Mésange bleue a pu être cartographiée (Figure 16), avec un R^2 égale à 0,18, ce qui ne permet cependant pas de la catégoriser dans les cartes représentatives de la réalité de terrain.

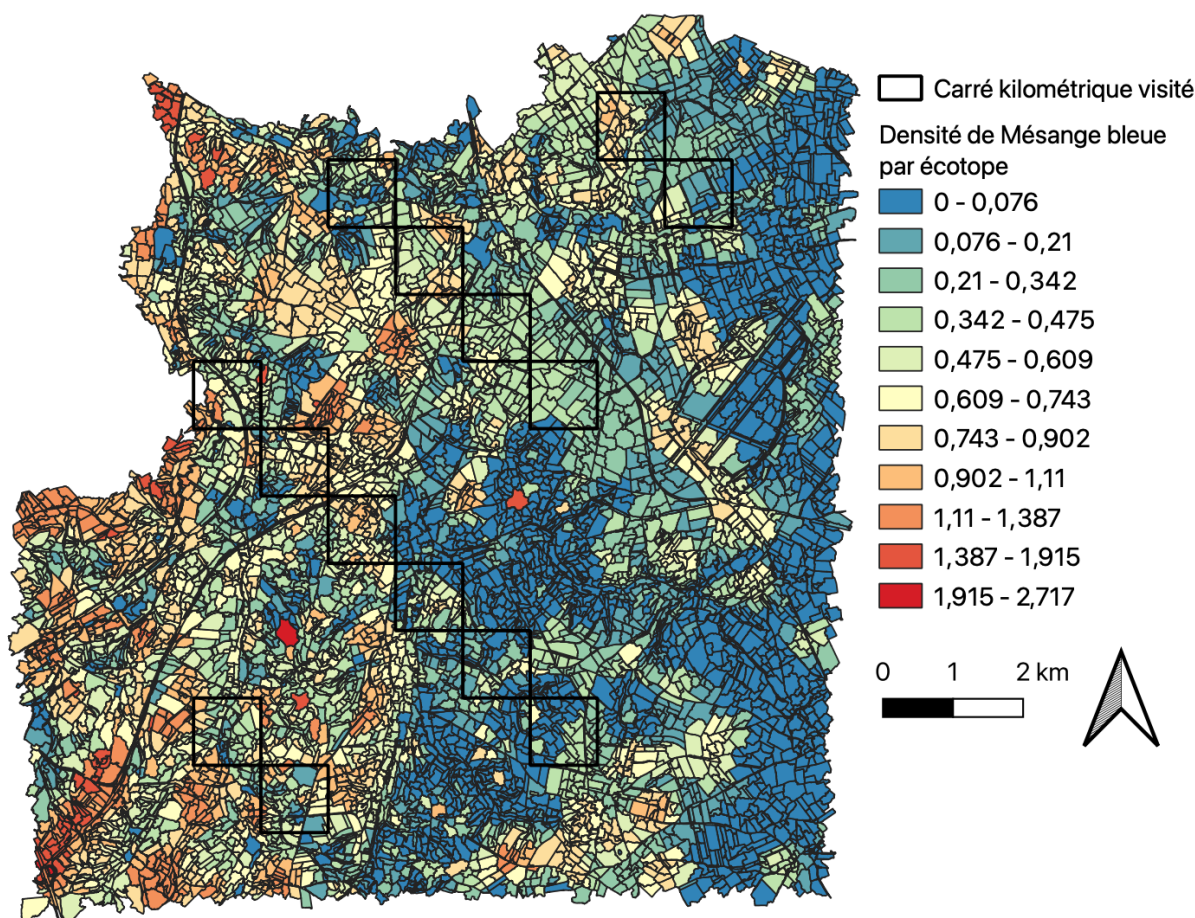


FIGURE 16 – Répartition de l’abondance de la Mésange bleue par écotope par une distribution ”Quasi-Poisson”.

16.5 Caractérisation d’une zone de territoire par groupe d’individus

Regrouper des polygones conjoints, dont l’abondance estimée pour une espèce est identique, réduit de manière significative la densité des espèces au kilomètre carré induisant une augmentation du R^2 (Tableau 11).

Au final, malgré l’augmentation significative des différents R^2 , cette nouvelle méthode de validation montre que le modèle n’a pas estimé correctement l’abondance de certaines espèces, telles que le Moineau domestique, le Pigeon ramier et la Tourterelle turque, dont le R^2 reste inférieure à 0,25. Cependant, d’autres espèces préalablement catégorisées comme mal modélisées, voient leur R^2 dépasser ce seuil. C’est le cas de la Corneille noire qui, en utilisant la méthode des zones territoriales, est classifiée comme une carte représentative de la réalité de terrain.

Concernant l'Alouette des champs, il est à noter que la nouvelle méthode de validation a été testée sur le modèle intégrant les données d'absence. La validation initiale du modèle (sans données d'absence) de cette espèce à une valeur de R^2 égale à 0,23 ; celui-ci augmente ensuite à 0,32 en intégrant les données d'absence et en utilisant toujours la première méthode de validation ; enfin, le R^2 atteint une valeur de 0,37 par la nouvelle méthode de validation sur le modèle intégrant les données d'absence.

TABLEAU 11 – Comparaison entre les méthodes de validation. La validation initiale somme l'abondance estimée par écotope, alors que la validation finale somme l'abondance estimée par zone de territoire définie par carré kilométrique.

Nom de l'espèce	Validation initiale (R^2)	Validation finale (R^2)
M. domestique (<i>Passer domesticus</i>)	0,07	0,18
P. ramier (<i>Columba palumbus</i>)	0,07	0,14
T. turque (<i>Streptopelia decaocto</i>)	0,08	0,12
C. noire (<i>Corvus corone</i>)	0,15	0,28
C. des tours (<i>Coloeus monedula</i>)	0,04	0,08
A. des champs (<i>Alauda arvensis</i>)	0,23	0,37
M. charbonnière (<i>Parus major</i>)	0,28	0,47
P. des arbres (<i>Fringilla coelebs</i>)	0,25	0,49

Discussion

17 Qualité des données issues de l'effort collectif et de terrain

Depuis le début de l'année 2022, 28 225 individus ont été observés, représentant 154 espèces d'oiseaux recensées au total sur le site d'étude, appartenant à 41 familles différentes. Ce recensement a été possible grâce à la participation de 141 utilisateurs sur le site Observations.be.

La qualité des données disponibles sur la plateforme paraît relativement bonne puisque seulement 7 % des observations sont notées incertaines, signifiant qu'elles sont en attente de validation mais pas fausses pour autant. Sur l'ensemble des données, un seul recensement a été refusé, ce qui semble indiquer que les observateurs ont les connaissances nécessaires pour reconnaître une espèce. Par ailleurs, les espèces étudiées étant généralement communes et non rares, les utilisateurs sont le plus souvent aptes à les identifier correctement. De plus, il existe un module d'intégration des données au niveau européen, le « Pan European Common Bird Monitoring Scheme », qui valide les résultats wallons à chaque transmission de données. Ceci augmente donc la qualité des résultats (DEROUAUX & PAQUET, 2018).

Lors des relevés de terrain réalisés pour cette étude entre février et mai 2022, 2289 individus ont été recensés, soit 47 espèces différentes et 28 familles. La qualité des données semble bonne puisque seulement 3,6 % des données sont catégorisées incertaines et aucune d'entre elles n'a été rejetée malgré le risque d'erreurs important (observateur non professionnel, difficulté de recenser toutes les espèces seul, choix des moments d'observations, etc.).

Bien que la campagne de terrain ait suivi un protocole standardisé, celle-ci ne s'est pas réalisée durant la meilleure période. En effet, elle a débuté vers le début du mois de mars et s'est terminée à la fin du mois d'avril. Cette période n'est pas la plus propice au recensement d'oiseaux. Il aurait fallu que la campagne s'étende au moins jusqu'à la fin mai / mi-juin. Le retour des températures estivales sont signes d'un retour des premiers migrants. Dès lors, des espèces, comme la Fauvette à tête noire et le Pouillot véloce, auraient pu être plus observées (PAUL GAILLY, s. d.). Comme la campagne a débuté tôt dans l'année, ces espèces migratrices n'ont pu être recensées qu'au deuxième passage par carré kilométrique, impliquant ainsi un manque de données. Dans le cas de cette étude, un troisième passage par carré aurait pu être bénéfique afin de collecter un maximum de données permettant alors une meilleure calibration du modèle de prédiction.

18 Répartition spatio-temporelle des données

De moins en moins d'individus ont été observés⁸ entre le 1^{er} janvier et la fin du mois de mai 2022 (cfr Section 11). Ceci peut s'expliquer par plusieurs raisons :

1. La campagne Natagora "Devine qui vient manger au jardin" s'est déroulée les 29 et 30 janvier 2022. Ayant attiré plus de 20 000 participants, il n'est pas surprenant que le nombre de recensements durant ce mois de janvier soit nettement supérieur aux autres mois de l'année ;
2. En hiver, les arbres étant dépourvus de feuilles, l'observation des oiseaux est rendue plus aisée ;
3. Les jours étant plus courts en hivers, l'activité des oiseaux commence plus tard, ce qui permet à plus d'observateurs d'être présents lors du pic d'activité ;
4. Au retour du printemps, le débourrement rend la tâche plus compliquée pour les observateurs ayant peu d'expérience ;
5. Aucune données issues d'Observations n'a été comptabilisé au mois de mai car les données ont été téléchargées le 24 avril 2022.

Dès lors, la variation temporaire des observations issues d'Observations.be et des relevés de terrain réalisés pour ce travail peut être mise en relation avec le cycle saisonnier ainsi que le moment de réalisations des différentes campagnes de recensement.

En ce qui concerne la variation spatiale, les milieux prospectés sont, par nombre d'observations décroissant, le milieu urbain, le milieu forestier, le milieu humide et le milieu agricole. Cela pourrait s'expliquer par le fait que la densité d'observateurs présents en ville/village est plus importante que celle en forêt. Il y a donc un potentiel d'observateurs plus large en centre urbain. De plus, les milieux urbains pourraient attirer plus d'individus car ils seraient une source de nourriture et d'abris. En outre, certaines espèces décrites comme forestières peuvent facilement s'observer dans les parcs de centre-ville ou de jardins comme la Mésange charbonnière (*Parus major*), le Pinson des arbres (*Fringilla coelebs*) ou le Troglodyte mignon (*Troglodyte troglodyte*), ce qui augmente d'autant plus le nombre d'oiseaux observés dans le milieu urbain.

8. Il est à noter que le nombre d'oiseaux observés ne correspond pas au nombre d'oiseaux réellement présents. Dès lors, le nombre d'oiseaux dans la nature n'est pas forcément décroissant au fil de l'année.

19 Modèle et méthode employée

19.1 Représentativité de l'échantillonnage

La représentativité spatiale de l'échantillonnage est assurée par une distribution diagonale régulière des carrés kilométriques sur l'ensemble du territoire wallon. Comme mentionné dans une section précédente (cfr Section 2.5.1), une couverture territoriale de l'ordre de 5 % (1600 carrés) est déjà suffisante pour modéliser la répartition des espèces peu ou moyennement répandues. Ceci permet de gagner un temps conséquent, car il n'est plus nécessaire de réitérer le suivi sur l'ensemble des 2800 carrés initialement prévus lors de la campagne de 2001-2007. L'ensemble des 14 carrés kilométriques, soit 14 km², représente 10,5 % de la superficie totale de la zone d'étude qui s'étend sur 132,74 kilomètres, représentant donc une surface suffisante pour assurer une bonne représentativité du territoire.

Cependant, l'observateur, qu'il soit professionnel ou non, choisit généralement des parcours qu'il connaît, des endroits qu'il apprécie particulièrement et où il peut retourner facilement d'une année à l'autre. Certains carrés kilométriques sont donc parfois préférés des observateurs, délaissant les autres. Ceci biaise les données territoriales, les carrés n'étant alors plus représentatifs car plus recensés de manière aléatoire.

En outre, la répétabilité interannuelle est un point de biais supplémentaire. Les observations peuvent fluctuer d'année en années suite à diverses pressions externes telles que les changements climatiques, les modifications d'occupation du sol, la pollution, etc.

Il faut également prendre en compte « l'effet observateur » (VORISEK, 2008). Cet effet se définit par la capacité de détection propre à chaque observateur : deux observateurs parcourant le même chemin au même endroit à la même heure n'auraient pas une prise de donnée identique.

De plus, les données issues des sciences participatives sont plus imprécises que celles issues de sciences professionnelles. Or, plus de la moitié des données recensées proviennent de source amatrice, ce qui indique que les données peuvent être grandement biaisées.

Ensuite, en moyenne, plus d'une espèce sur huit recensées est une Mésange bleue ou charbonnière, ce qui constitue une autre biais. En effet, ces deux espèces fréquentent régulièrement les mangeoires et sont performantes pour repérer une source de nourriture si elle est présente (VIGIE-NATURE, s. d.).

L'utilisation de sciences participatives implique l'existence d'un décalage entre la biodiversité réelle et la biodiversité perçue. De fait, les observateurs novices n'ont aucune obligation d'inventorier toutes les espèces différentes et leur nombre exact lors de leurs visites ou balades ("l'effet observateur"). Cela est d'autant plus vrai que des opérations spécifiques ne ciblent que certaines espèces ou ordres, par exemple les Passereaux venant se nourrir aux mangeoires ou les oiseaux d'eau pour les Dénombrements Hivernaux d'Oiseaux d'Eau (DHOE) (AVES, 2022). Néanmoins, réduire les coûts de prospection (main-d'œuvre, temps) et permettre l'exploration de lieux privés et difficilement accessibles sont des avantages à ne pas négliger des sciences participatives. (VIGIE-NATURE, s. d.).

19.2 Précision spatiale des données

Tout d'abord, la précision des données est directement corrélée à la manière dont celles-ci sont encodées (papier ou application mobile, en mouvement ou à l'arrêt) ainsi qu'au moyen d'observation (ouïe, visuel). Tout d'abord, il existe deux moyens de prendre note des individus observés sur le terrain : le premier étant de compléter une fiche de terrain et le deuxième étant d'enregistrer les observations sur une application mobile. La fiche est complétée sur le terrain et ensuite l'observateur enregistre ses observations sur la plateforme Observations.be. Les applications mobiles, quant à elle, permettent d'encoder directement les observations. Elles fournissent également l'avantage de localiser directement les individus observés via le GPS de l'appareil encodant. Dès lors, l'observateur n'encode plus sa position, mais peut situer immédiatement d'où l'espèce a été vue/entendue. Cependant, la précision étant liée au GPS de l'appareil, cette-dernière accuse une marge d'erreur plus importante lorsque l'observateur est en mouvement.

Par ailleurs, il est également possible de percevoir un individu de deux façons. Visuellement, il est aisé d'estimer correctement la distance qui sépare l'observateur de l'individu. Par contre, il est plus compliqué de l'évaluer lorsque l'individu est entendu sans contact visuel. En effet, certaines espèces ont des chants très portant qui peuvent s'entendre à plusieurs dizaines de mètres de distance de la source, par exemple le tambourinage des pics (NAGUIB, 1996).

Comme abordé précédemment, les données issues du site Observation.be ont une variabilité de précision spatiale assez importante, allant de 1 à 999 mètres. Certains observateurs remplissent leur encodage sur une fiche de terrain une fois rentrés chez eux. Pour cette raison, une précision spatiale de 999 mètres est présente dans le jeu de données. Cependant, cela ne représente que 4 observations sur l'ensemble des données. En terme général, la précision moyenne approche les 14 mètres, ce qui reste très acceptable. Concernant la précision spatiale des relevés de terrain, celle-ci varie de 5 à 60 mètres et la moyenne avoisine les 11 mètres.

L'imprécision des données peut avoir des conséquences non négligeables sur les résultats de notre modèle. En effet, au moment de géolocaliser l'observation, cette erreur de précision peut associer l'oiseau à un écotope différent de celui où l'individu est réellement présent. Ainsi, au moment d'associer à l'oiseau les variables environnementales (liées à l'écotope dans lequel il a été "soi-disant" observé), des potentielles erreurs d'estimation peuvent être commises par le modèle.

Afin d'éviter cela, un rayon de 25 mètres de précision pour la localisation des individus a été choisi. Celui-ci est cohérent avec la précision moyenne des observations. De fait, plus de 95 % des observations issues d'Observation.be et des données de terrain comprises ont une précision comprise entre 0 et 25 mètres. Les 5 % restant ont été supprimés pour éviter les erreurs de calibration du modèle.

Dans l'optique d'améliorer encore plus la précision spatiale, il serait préférable que chaque utilisateur utilise les applications mobiles telles que iObs ou Obsmaps. Ceci permettrait de rendre ces données plus faciles d'utilisation et améliorer leur qualité. Ces applications offrent divers avantages tant pour les utilisateurs que pour les gestionnaires de données. Comme mentionné dans une des sections précédentes (cfr Section 7.2), elles permettent d'encoder la position géographique, la date et l'heure. L'utilisateur peut dès lors encoder directement, de manière précise, la localisation de l'espèce observée augmentant considérablement la qualité du jeu de données puisque la position sera celle de l'observation et non plus celle de l'utilisateur. De plus, la précision sera d'autant plus fine puisque l'observation sera enregistrée directement sur place.

Lors des prochaines campagnes de recensement, il serait conseillé de s'affranchir des fiches de terrain afin de se digitaliser.

19.3 Polygones et écotopes

Les problèmes liés à l'hétérogénéité des unités d'échantillonnage de polygones sont similaires aux problèmes dits des pixels mixtes. Les règles d'étiquetage doivent impérativement être définies clairement afin d'éviter des résultats de validation ambigus. Sur base de cet étiquetage, un paysage peut être décrit avec différentes légendes comme illustré à la Figure 17. L'étendue de la zone urbaine est largement affectée en fonction du type de légende choisi (JULIEN RADOUX & PATRICK BOGAERT, 2017).

En fonction de la classification choisie, une espèce pourrait se voir être associée à deux milieux différents.

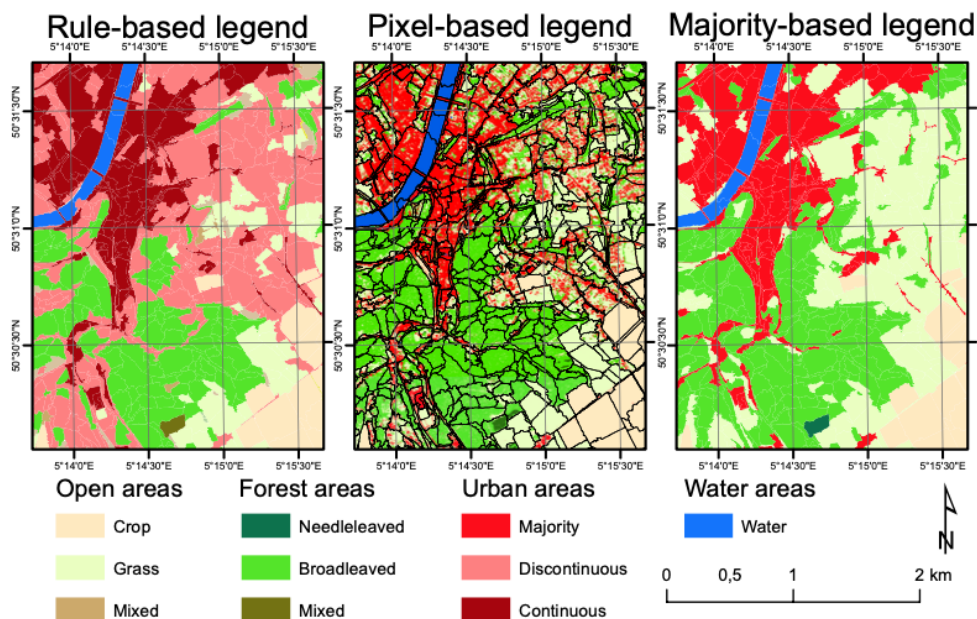


FIGURE 17 – Différentes légendes appliquées sur le même paysage. L'image centrale est une représentation basée sur les pixels avec une légende de la règle de la majorité. Deux autres représentations sont construites en cohérence avec cette classification. À gauche, les classes sont basées sur un ensemble de seuils dérivés de règles de type LCCS. À droite, une légende basée sur la majorité des cellules est utilisée (JULIEN RADOUX & PATRICK BOGAERT, 2017).

Cependant, cette classification se base sur les différentes composantes que constituent chaque écotope. Ces composantes ont été définies par photographies aériennes et données LIDAR et fournissent des informations pour la segmentation du paysage en fonction de la signature spectrale, de la structure en hauteur et de la topographie (RADOUX et al., 2019). Cependant, les variables de certains écotopes ne sont pas toujours définies correctement. En effet, la somme des différents pourcentages d'occupation du sol d'un écotope doit être égale à 100 %. Or, plusieurs écotopes n'atteignent pas ces 100 %. C'est pour cette raison qu'une variable "Autre" a été créée pour combler le déficit. Il arrive même que ce pourcentage résiduel d'occupation du sol soit prédominant parmi les variables. Ceci nuit à la prédiction de l'abondance des espèces par le modèle. C'est le cas de l'Alouette des champs (cfr Section 16.2) et de la Tourterelle turque (Annexe O) et de, où le modèle prédit de manière abusive une grande présence de chaque espèce au sein de 5 écotopes. Après vérification, ces 5 écotopes ont une variable "Autre" valant plus de 95 %.

19.4 Le modèle

Sur les 41 espèces recensées dans cette étude, le modèle a pu identifier un jeu de variables explicatives pour 9 d'entre-elles. Parmi ces 9 espèces, deux ont pu être modélisées convena-

blement selon la clé dichotomique. Il s'agit de la Mésange charbonnière ($R^2 = 0,28$) et du Pinson des arbres ($R^2 = 0,25$).

Concernant les sept espèces liées au milieu aquatique et aux plans d'eau précisément, le manque de données ne permet pas au modèle d'identifier un jeu de variables explicatives. Néanmoins, le milieu humide n'a été que faiblement représenté dans ce travail puisqu'un seul étang peu étendu a été inventorié au sein des 14 carrés kilométriques. Par conséquent, ce faible nombre de sites recensés ne permet pas d'obtenir un modèle correctement calibré pour ces différentes espèces. Ainsi, le modèle détermine uniquement l'intercept. Dès lors, une extrapolation pourrait prédire les abondances des espèces du milieu humide. Par exemple, si un Colvert a été observé dans un étang de 10 m², une centaine de Colverts devrait être observée sur un plan d'eau d'une superficie de 100 m². Il faut cependant rester prudent lors d'extrapolation car tous les plans d'eau ne sont pas comparables en termes de qualité des eaux, de profondeur et de type de végétations, mais aussi en termes d'espèces présentes et d'interactions possibles.

Afin d'évaluer l'impact dans la modification de l'utilisation des données, la manipulation de certaines variables, le choix d'un modèle différent et la méthode d'agrégation, plusieurs tests de sensibilité du modèle ont été réalisés et sont critiqués ci-dessous.

19.4.1 Utilisation des données issues de la campagne de terrain uniquement

Pour rappel, il serait intéressant d'utiliser notre modèle avec seulement des données de comptage puisque les données de présence pourraient entraîner des biais dans l'estimation d'abondance des espèces par écotope. Les données issues des sciences participatives ne suivent pas un protocole officiel et ne garantissent pas un apport de données de comptage, au contraire des données de terrain.

Dans cette optique, les données de terrain de la Mésange charbonnière ont été sélectionnées pour vérifier ce test, car il s'agit de l'espèce la plus observée (123 individus enregistrés). Malheureusement, le modèle n'a pas pu déterminer un jeu de variables explicatives permettant d'obtenir une cartographie, ce qui reste cohérent avec la classification de la clé dichotomique.

Dès lors, il n'a pas été possible de comparer la qualité des résultats entre un modèle utilisant tout type de données confondues et un modèle utilisant uniquement des données de comptage. La prospection d'un nombre de carrés kilométriques plus important permettrait d'acquérir un plus grand nombre de données afin de générer un résultat. Ce manque de données a donc été comblé par les données issues du site Observations.be, ce qui démontre toute l'importance

de celles-ci. En effet, même si elles sont d’une qualité possiblement moindre, elles ont tout de même permis d’obtenir des résultats concluant pour deux espèces (la M. charbonnière et le P. des arbres).

Finalement, il faudrait comparer la valeur de validation d’un modèle uniquement basé sur des données de comptages à celle de notre modèle afin de déterminer si le premier offre un meilleur résultat que le second.

19.4.2 Ajout des données d’absence

L’ajout des données d’absence a un impact positif non négligeable. En effet, pour le modèle de l’Alouette des champs, la validation est passée d’un R^2 égal à 0,23 à un R^2 égal à 0,32. À l’avenir, lors de l’utilisation d’un modèle prédictif avec une régression de Poisson, l’utilisation de données d’absence pourraient donc être indispensable dans certains cas. D’autres validations de modèle devrait être testées afin de confirmer cette affirmation.

Il est à noter que les données d’absence ont pu être déterminées grâce aux données issues d’Observations.be. Ce test montre donc, de surcroît, l’intérêt des données apportées par les sciences participatives.

De plus, grâce à l’intégration de ces données d’absence, la nouvelle cartographie de l’abondance de l’A. des champs est estimée représentative de la réalité de terrain pour un nombre d’individus recensé inférieur au seuil minimum. L’ajout de ces données d’absence permettrait alors d’obtenir des résultats corrects avec un plus faible nombre d’individus enregistrés. Dès lors, le seuil minimal nécessaire d’observations pourrait être revu à la baisse.

19.4.3 Dépendance de la variable de référence

Pour rappel, la variable Plowed250 a été choisie comme variable de référence pour pallier les problèmes de colinéarité entre variables. Ce choix a été fait car cette variable est la plus répandue et représente correctement la matrice d’occupation du sol de la zone d’étude.

Cette étude a démontré que le choix de cette variable de référence a des répercussions sur la qualité des résultats. Précisément, lorsque la variable de référence utilisée est Plowed250, la validation du modèle obtient un R^2 de 0,25 catégorisant le résultat comme une cartographie représentative de la réalité de terrain. Lorsqu’une autre variable que Plowed250 est employée, la validation des cartographies est inférieure à 0,25 (cfr Section 16.3).

Afin de s’assurer que le choix de Plowed250 est le plus judicieux, toutes les variables d’occupation du sol à 250 mètre de rayon de l’écotope devraient être testées comme variables de référence.

19.4.4 Utilisation d'une régression de type Quasi-Poisson

Les suppositions majeures des PRM sont (1) que les données suivent une distribution de Poisson et (2) que les réponses sont indépendantes. Cette deuxième condition signifie que l'observation d'une espèce n'affecte pas la probabilité d'en apercevoir une autre au même endroit. Cependant, ce dernier point est rarement respecté dans la nature. Le non-respect de ces suppositions peut mener à une dispersion des observations. Cela signifie qu'une différence entre l'espérance et la variance est observée due soit à une observation de valeurs extrêmes plus souvent que prévu, soit à une répartition régulière des observations. Ces problèmes potentiels ont mené à développer plusieurs extensions de PRM telles que, par exemple, *zero inflated Poisson regression*, *truncated Poisson regression* (J. MICHAEL SCOTT et al., 2001), ainsi que *negative binomial regression* et *Quasi-Poisson regression* (VER HOEF & BOVENG, 2007).

La régression Quasi-Poisson peut être utilisée dans le cas de données sur- ou sousdispersées. Autrement dit, la régression Quasi-Poisson est utilisée dans le cas où le coefficient de variation n'est pas égal à 1. La dispersion de données représente tout ensemble de données ou modèle où la variance est différente de la moyenne (VER HOEF & BOVENG, 2007). En écologie, les données de comptage sont généralement sujettes à une surdispersion, présentant une plus grande variabilité que prévu dans un modèle statistique donné (MACKENZIE J. EDMONDSON et al., 2022). Un moyen de faire face à cette dispersion des données est d'utiliser une variante des modèles linéaires généralisés (P. MCCULLAGH & J. A. NELDER, 1989), où l'approche la plus commune est le modèle "Quasi-Poisson" (WEDDERBURN, 1974).

Pour la Mésange bleue, le modèle de Poisson ne permettait pas de sortir un ensemble de variable par AIC car l'effectif enregistré était trop faible. Le modèle n'a alors pas pu définir un jeu de variables explicatives. En appliquant le modèle Quasi-Poisson, un jeu de variables a pu être déterminé, ainsi que le paramètre de dispersion. Celui-ci vaut 0,4 impliquant un cas de sousdispersion. Avec ce modèle, le résultat de l'abondance de la Mésange bleue n'est pas idéal ($R^2 = 0,18$) mais permettrait de se rapprocher d'un résultat concluant si d'autres améliorations étaient apportées au modèle. Il est également à rappeler que les variables explicatives n'ont pas été sélectionnées de façon automatique et qu'une sélection informatisée pourrait définir un jeu de variables plus significatif pour modéliser l'abondance de l'espèce.

En conclusion, en changeant de modèle, plusieurs autres espèces caractérisées par un faible nombre d'individus recensés pourraient être modélisées.

19.4.5 Méthode d'agrégation

Deux méthodes d'agrégation ont été présentées et développées dans ce rapport.

La première méthode consiste à sommer les abondances, calibrées en fonction de la surface de l'écotope, au sein des carrés kilométriques. On observe qu'en utilisant cette méthode, l'abondance par kilomètre carré des espèces est bien plus élevée que l'abondance estimée en 2015, avec parfois des abondances jusqu'à vingt fois plus importantes. Une hypothèse est avancée afin d'expliquer ce phénomène :

Lorsque deux individus sont observés au sein d'un écotope, les variables environnementales associées à celui-ci leur sont associées. Le modèle prédictif estime alors le même nombre d'individus pour tous les écotopes présentant des variables environnementales similaires. Or, ces écotopes sont souvent conjoints. Au final, même si seulement deux individus ont été observés dans un seul écotope, le modèle considère que tous les écotopes similaires adjacents contiennent également deux individus chacun. Ceci a pour conséquence d'augmenter considérablement la densité de l'espèce observée lors de l'agrégation, impactant la valeur du R^2 .

La deuxième méthode consiste à supposer que, lorsque le modèle prédit des abondances identiques pour des écotopes limitrophes, ces écotopes représentent en réalité une zone territoriale du groupe d'individu observé au sein d'un de ces écotopes. En groupant alors ceux-ci, et en sommant ensuite les abondances estimées par zones de territoire défini, celles-ci se rapprochent des abondances mesurées en 2015, augmentant la validation des cartographies.

Cette méthode, même si elle permet d'augmenter le R^2 de toutes les espèces, reste très systématique. En effet, il se pourrait très bien que deux groupes d'individus différents soient présents dans 2 écotopes adjacents. Or, la méthode d'agrégation proposée ici ne prendrait pas en compte les deux groupes d'individus. Une méthode aussi systématique reste donc imprécise et ne semble dès lors pas pouvoir générer de R^2 très élevé.

En outre, cette méthode a été réfléchi pour augmenter les R^2 . Cependant, l'objectif de la méthode est de mesurer la validation réelle des résultats et non d'orienter les calculs pour que celle-ci soit la plus élevée possible.

Il serait donc bénéfique pour les futures études cartographiques de déterminer une méthodologie de validation des résultats la plus juste possible, afin de les homogénéiser.

20 Perspectives d'étude

Le fait de connaître l'abondance des différentes espèces aviaires à l'échelle des écotopes plutôt qu'à l'échelle des carrés kilométriques a permis d'obtenir des cartographies plus précises. Ces différentes cartes pourraient dès lors servir d'outil d'aide à la décision pour améliorer les plans de gestion associés à la préservation de l'avifaune. Par exemple, connaître précisément où se localise une espèce permettrait de connaître exactement l'impact qu'une modification d'occupation du sol suite à une cause humaine ou naturelle pourrait avoir sur la population de cette espèce. Ceci est d'autant plus vrai pour les espèces aux exigences spécifiques telles que l'Alouette lulu (*Lullula arborea*), endémique aux pelouses sèches calcicoles (GERBER et al., 2006), qui sont dès lors plus sensibles à ces perturbations.

Pour cette étude, la résolution spatiale basée sur les écotopes a été choisie, ce qui correspond en moyenne à un hectare de résolution. Pour aller plus loin, connaître la spécificité du territoire permettrait d'obtenir des résultats cartographiques plus précis. L'influence des petits éléments structuraux du paysage (arbres, haies, mares, ...), ainsi que l'échelle de résolution d'étude des relations entre espèces et habitat font déjà l'objet de beaucoup d'études (cfr Section 2.7.2). Cartographier l'abondance des oiseaux à plus haute résolution spatiale que les écotopes, en prenant en compte les éléments structuraux du paysage, pourrait dès lors être intéressant à approfondir.

Il faut également prendre en compte que certaines espèces ont des exigences écologiques moins strictes, elles sont plus communément appelées des espèces « généralistes ». Elles sont donc présentes dans une plus large gamme d'habitats. Il est à supposer que ces espèces ne nécessiteraient probablement pas une si grande précision dans la résolution spatiale puisqu'elles sont, par définition, présentes partout dans le territoire. Pour les futures études, il serait donc pertinent de définir la résolution d'étude optimale pour chaque espèce afin que l'effort de recherche soit le plus rentable possible.

Conclusion

Ce travail a permis de mettre en évidence la possibilité de passer d'une cartographie de l'abondance de populations d'oiseaux à faible résolution à une cartographie à une résolution plus élevée. À l'heure actuelle, le modèle adopté par Aves a une résolution spatiale d'un kilomètre carré. Le modèle utilisé dans cette étude, suivant une régression de Poisson, a permis de prédire l'abondance des espèces à l'échelle des écotopes, soit une réduction de résolution spatiale d'un facteur cent. La difficulté principale de cette étude a été de récolter suffisamment de données lors de la campagne de terrain. Il n'a dès lors pas été possible d'obtenir des résultats cartographiques pour toutes les espèces recensées.

Par la même occasion, ce rapport a mis en valeur l'utilisation des sciences participatives et des plateformes d'encodage en ligne qui se sont révélées de bonne qualité et ont aidé à obtenir les résultats escomptés.

À l'avenir, une formation des contributeurs, concernant les méthodes d'encodage de données et les protocoles, permettrait de perfectionner davantage la qualité de ces données.

Finalement, il est possible d'obtenir des résultats adéquats, à condition de disposer de suffisamment de données et d'utiliser une méthode d'analyse appropriée.

Cette étude a également illustré que la façon dont le modèle est codifié et l'utilisation de données du type "donnée d'absence" ont des impacts importants sur les résultats, qu'ils soient positifs ou négatifs. L'emploi d'une régression du type Quasi-Poisson peut également être une solution lorsqu'il faut faire face à des données dispersées. Ou encore, il faudrait déterminer une méthode de validation appropriée.

En conclusion, ce travail a démontré qu'un effort additionnel restreint lors des campagnes de terrain, apporte une réelle plus-value aux résultats. En effet, renseigner directement la position géographique des observations est une amélioration qui permet d'obtenir de meilleurs résultats. Ainsi, les plans de gestion sur l'avifaune pourront s'appuyer sur des cartographies précises et apporter les meilleures solutions possibles. Dans la même optique, un autre point qui pourrait améliorer les plans de gestion de l'avifaune serait de caractériser l'abondance des espèces à plus haute résolution, c'est-à-dire en prenant en compte les éléments du paysage. De plus amples études restent à réaliser pour confirmer ce point.

Références

- AEBISCHER, N.J., GREEN, R.E & EVANS, A.D. (2000). From science to recovery : four case studies of how research has been translated into conservation action in the UK.
- AIZPURUA, O., PAQUET, J.-Y., BROTONS, L. & TITEUX, N. (2015). Optimising long-term monitoring projects for species distribution modelling : how atlas data may help. *Ecography*, 38(1), 29-40. <https://doi.org/10.1111/ecog.00749>
- ANTOINE DEROUAUX. (2021). Suivi des oiseaux communs en Wallonie 1990-2020. Récupérée 16 mars 2022, à partir de https://drive.google.com/file/d/1O2AsomcU0wGk0q4gCrm-HwmLZBcrOIxf/view?usp=embed_facebook
- AVES. (2010). *Le programme de Surveillance des Oiseaux Communs en Wallonie SOCWAL. Notice à l'intention des collaborateurs - saison 2010.* (rapp. tech.).
- AVES. (2015). Echantillonnage 2015-2017 des oiseaux nicheurs en Wallonie : notice à l'attention des observateurs.
- AVES. (2022). Les dénombrements hivernaux d'oiseaux d'eau (DHOE) - Aves. Récupérée 2 juin 2022, à partir de <https://aves.natagora.be/coa/rhoe>
- Aves, pôle ornithologique de Natagora, association de protection de la nature. (s. d.). Récupérée 12 novembre 2021, à partir de <https://www.aves.be/index.php?id=aves.asbl>
- BAILLIE, S. R. & SCHAUB, M. (2009). Understanding changes in bird populations – the role of bird marking. *Ringling & Migration*, 24(3), 189-198. <https://doi.org/10.1080/03078698.2009.9674391>
- BALDERAMA, E., GARDNER, B. & REICH, B. J. (2016). A spatial-temporal double-hurdle model for extremely over-dispersed avian count data. *Spatial Statistics*, 18, 263-275. <https://doi.org/10.1016/j.spasta.2016.05.001>
- BARBU, C., PAPAIX, J., FABRE, F., LAVIGNE, C., ALIGNIER, A., ERNOULT, A. & POGGI, S. (2019). *Analyses statistiques des relations entre mosaïque paysagère et processus écologiques.*
- BESBEAS, P., FREEMAN, S. N., MORGAN, B. J. T. & CATCHPOLE, E. A. (2002). Integrating Mark-Recapture-Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters [Publisher : [Wiley, International Biometric Society]]. *Biometrics*, 58(3), 540-547. Récupérée 15 mars 2022, à partir de <https://www.jstor.org/stable/3068577>
- BIAN, L. (2007). Object-Oriented Representation of Environmental Phenomena : Is Everything Best Represented as an Object? *Annals of the Association of American Geographers*, 97(2), 267-281. <https://doi.org/10.1111/j.1467-8306.2007.00535.x>
- BIBBY, C. J., BURGESS, N. D., HILLIS, D. M., HILL, D. A. & MUSTOE, S. (2000). *Bird Census Techniques* [Google-Books-ID : Ld5wkzPp49cC]. Elsevier.

- BIRDLIFE. (2013). BirdLife Data Zone. Récupérée 16 décembre 2021, à partir de <http://datazone.birdlife.org/sowb/casestudy/birds-are-very-useful-indicators-for-other-kinds-of-biodiversity>
- BLOCK, W. M. & BRENNAN, L. A. (1993). The Habitat Concept in Ornithology : Theory and Applications. In D. M. POWER (Éd.), *Current Ornithology* (p. 35-91). Springer US. https://doi.org/10.1007/978-1-4757-9912-5_2
- BURNS, F., EATON, M. A., BURFIELD, I. J., KLVAŇOVÁ, A., ŠILAROVÁ, E., STANEVA, A. & GREGORY, R. D. (2021). Abundance decline in the avifauna of the European Union reveals cross-continental similarities in biodiversity change [eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ece3.8282>]. *Ecology and Evolution*, n/a(n/a). <https://doi.org/10.1002/ece3.8282>
- CARO, T. M. & O'DOHERTY, G. (1999). On the Use of Surrogate Species in Conservation Biology. *Conservation Biology*, 13(4), 805-814. <https://doi.org/10.1046/j.1523-1739.1999.98338.x>
- CASTILLA, G., HERNANDO, A., ZHANG, C., MAZUMDAR, D. & McDERMID, G. J. (2012). AN INTEGRATED FRAMEWORK FOR ASSESSING THE ACCURACY OF GEOBIA LANDCOVER PRODUCTS, 4.
- CAVANAUGH, J. E. & NEATH, A. A. (2019). The Akaike information criterion : Background, derivation, properties, application, interpretation, and refinements. *WIREs Computational Statistics*, 11(3). <https://doi.org/10.1002/wics.1460>
- CONSUL, P. & FAMOYE, F. (1992). Generalized poisson regression model. *Communications in Statistics - Theory and Methods*, 21(1), 89-109. <https://doi.org/10.1080/03610929208830766>
- COUCLELIS, H. (1996). Towards an Operational Typology of Geographic Entities with Ill-defined Boundaries [Num Pages : 11]. *Geographic Objects with Indeterminate Boundaries*. CRC Press.
- DELANGRE, J., RADOUX, J. & DUFRÊNE, M. (2018). Landscape delineation strategy and size of mapping units impact the performance of habitat suitability models. *Ecological Informatics*, 47, 55-60. <https://doi.org/10.1016/j.ecoinf.2017.08.005>
- DEROUAUX & PAQUET. (2018). L'évoLution préoccupante des popuLations d'oiseaux nicheurs en WaLLonie : 28 ans de surveiLLance de L'avifaune commune.
- DURO, D. C., FRANKLIN, S. E. & DUBÉ, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment*, 118, 259-272. <https://doi.org/10.1016/j.rse.2011.11.020>

- FLETCHER, R. J., HEFLEY, T. J., ROBERTSON, E. P., ZUCKERBERG, B., MCCLEERY, R. A. & DORAZIO, R. M. (2019). A practical guide for combining data to model species distributions. *Ecology*, e02710. <https://doi.org/10.1002/ecy.2710>
- FOUCART, T. (2006). Colinéarité et régression linéaire. *Mathématiques et sciences humaines*, (173). <https://doi.org/10.4000/msh.2963>
- GERBER, A., MÜLLER, S., SCHWALLER, T., SCHMID, H. & SPAAR, R. (2006). Répartition de l'Alouette lulu *Lullula arborea* dans le Jura suisse : situation actuelle et évolution depuis la fin des années 1970. *Nos Oiseaux*, 53, 131-144.
- GOTTSCHALK, T. K., AUE, B., HOTES, S. & EKSCHMITT, K. (2011). Influence of grain size on species-habitat models. *Ecological Modelling*, 222(18), 3403-3412. <https://doi.org/10.1016/j.ecolmodel.2011.07.008>
- GREGORY, R. D. (2000). DEVELOPMENT OF BREEDING BIRD MONITORING IN THE UNITED KINGDOM AND ADOPTING ITS PRINCIPLES ELSEWHERE. *THE RING*, 10.
- GREGORY, R. D. & STRIEN, A. v. (2010). Wild Bird Indicators : Using Composite Population Trends of Birds as Measures of Environmental Health. *Ornithological Science*, 9(1), 3-22. <https://doi.org/10.2326/osj.9.3>
- GREGORY ET AL., R. D. (2005). Strien A van, Vorisek P, Meyling AWG, Noble DG, Foppen RPB, et al. 360, 269-288.
- GRINDE, A. R., NIEMI, G. J., STURTEVANT, B. R., PANCI, H., THOGMARTIN, W. & WOLTER, P. (2017). Importance of scale, land cover, and weather on the abundance of bird species in a managed forest. *Forest Ecology and Management*, 405, 295-308. <https://doi.org/10.1016/j.foreco.2017.09.057>
- GUISAN, A., EDWARDS, T. C. & HASTIE, T. (2002). Generalized linear and generalized additive models in studies of species distributions : setting the scene. *Ecological Modelling*, 157(2-3), 89-100. [https://doi.org/10.1016/S0304-3800\(02\)00204-1](https://doi.org/10.1016/S0304-3800(02)00204-1)
- GUISAN, A. & ZIMMERMANN, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2-3), 147-186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- HABOU, K. M. A., ABDOU, L., RABIOU, H. & MAHAMANE, A. (2021). Impact des changements climatiques sur la dynamique de l'habitat potentiel de *Balanites aegyptiaca* (L.) Del. au Niger. *Revue Marocaine des Sciences Agronomiques et Vétérinaires*, 9(2). Récupérée 29 juillet 2022, à partir de https://www.agrimaroc.org/index.php/Actes_IAPH2/article/view/947
- HATTAB, T., BEN RAIS LASRAM, F. & SAMMARI, CH. (2011). Bulletin de l'Institut National des Sciences et Technologies de la Mer, 65-71.

- HILBE, J. M. (2011). *Negative binomial regression* (2nd ed) [OCLC : ocn694679188]. Cambridge University Press.
- IAN NEWTON. (1998). Population Limitation in Birds.
- J. MICHAEL SCOTT, HEGLUND, P. J. & MICHAEL L. MORRISON. (2001). *Predicting Species Occurrences : issues of accuracy and scale*. Island Press.
- JOHN L. HARPER & DAVID L. HAWKSWORTH. (1994). Biodiversity : measurement and estimation. Preface. *Philosophical Transactions of the Royal Society of London. Series B : Biological Sciences*, 345(1311), 5-12. <https://doi.org/10.1098/rstb.1994.0081>
- JONES, J. (2001). Habitat Selection Studies in Avian Ecology : A Critical Review. *The Auk*, 118(2), 557-562. <https://doi.org/10.1093/auk/118.2.557>
- JULIEN RADOUX & PATRICK BOGAERT. (2017). *Good Practices for Object-Based Accuracy Assessment* (rapp. tech.).
- JULLIARD, R., CLAVEL, J., DEVICTOR, V., JIGUET, F. & COUVET, D. (2006). Spatial segregation of specialists and generalists in bird communities. *Ecology Letters*, 9(11), 1237-1244. <https://doi.org/10.1111/j.1461-0248.2006.00977.x>
- KASSEN, R. (2002). The experimental evolution of specialists, generalists, and the maintenance of diversity : Experimental evolution in variable environments. *Journal of Evolutionary Biology*, 15(2), 173-190. <https://doi.org/10.1046/j.1420-9101.2002.00377.x>
- KUNIN, W. E. & GASTON, K. J. (1993). The biology of rarity : Patterns, causes and consequences. *Trends in Ecology & Evolution*, 8(8), 298-301. [https://doi.org/10.1016/0169-5347\(93\)90259-R](https://doi.org/10.1016/0169-5347(93)90259-R)
- La biodiversité en Wallonie. (s. d.). Récupérée 19 novembre 2021, à partir de <http://biodiversite.wallonie.be/fr/le-portail.html?IDC=3291>
- LAMBERTINI, M. (2020). *Living Planet Report 2020 : Bending the Curve of Biodiversity Loss* [OCLC : 1220988520]. Récupérée 15 novembre 2021, à partir de <https://www.zsl.org/sites/default/files/LPR%202020%20Full%20report.pdf>
- Les contributions volontaires en nature dans les associations [Section : Actualités]. (2018). Récupérée 1 juin 2022, à partir de <https://institut-isbl.fr/contributions-volontaires-nature-associations/>
- LIFEWATCH. (s. d.). Lifewatch regional portal. Récupérée 12 novembre 2021, à partir de <https://www.lifewatch.be/>
- LIN, Y.-P., LIN, W.-C. & WU, W.-Y. (2015). Uncertainty in Various Habitat Suitability Models and Its Impact on Habitat Suitability Estimates for Fish. *Water*, 7(12), 4088-4107. <https://doi.org/10.3390/w7084088>
- LINDENMAYER, D. B., GIBBONS, P., BOURKE, M., BURGMAN, M., DICKMAN, C. R., FERRIER, S., FITZSIMONS, J., FREUDENBERGER, D., GARNETT, S. T., GROVES, C., HOBBS, R. J., KINGSFORD, R. T., KREBS, C., LEGGE, S., LOWE, A. J., MCLEAN, R., MONTAMBAULT,

- J., POSSINGHAM, H., RADFORD, J., ... ZERGER, A. (2012). Improving biodiversity monitoring : IMPROVING BIODIVERSITY MONITORING. *Austral Ecology*, 37(3), 285-294. <https://doi.org/10.1111/j.1442-9993.2011.02314.x>
- MACKENZIE J. EDMONDSON, CHONGLIANG LUO, MD. NAZMUL ISLAM, NATALIE E. SHEILS, JOHN BURESH, ZHAOYI CHEN, JIANG BIAN & YONG CHEN. (2022). *Distributed Quasi-Poisson regression algorithm for modeling multi-site count outcomes in distributed data networks - ScienceDirect* (rapp. tech.). Récupérée 16 juillet 2022, à partir de <https://www.sciencedirect.com/science/article/abs/pii/S1532046422001137>
- MARSHALL, M., DIEFENBACH, D., WOOD, L. & COOPER, R. (2004). Annual survival estimation of migratory songbirds confounded by incomplete breeding site-fidelity : Study designs that may help. *Animal Biodiversity and Conservation*, 27.
- MI, C., HUETTMANN, F., GUO, Y., HAN, X. & WEN, L. (2017). Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ*, 5, e2849. <https://doi.org/10.7717/peerj.2849>
- MORRISON, M. L., MARCOT, B. & MANNAN, W. (2012). *Wildlife-Habitat Relationships : Concepts and Applications* [Google-Books-ID : yh2AwqA3mrYC]. Island Press.
- MOUKRIM, S., LAHSSINI, S., ALAOUI, H. M., RIFAI, N., ARAHOU, M. & RHAZI, L. (2018). Modélisation de la distribution spatiale des espèces endémiques pour la conservation : "cas de l'Arginia spinosa", 14.
- NAGUIB, M. (1996). Auditory distance estimation in song birds : Implications, methodologies and perspectives. *Behavioural Processes*, 38(2), 163-168. [https://doi.org/10.1016/0376-6357\(95\)00077-1](https://doi.org/10.1016/0376-6357(95)00077-1)
- Natagora [file : ///Users/memekaye/Downloads/Rapport_annuel_Natagora2020_BD.pdf]. (2018). Récupérée 12 novembre 2021, à partir de <https://www.natagora.be/qui-est-natagora>
- OBSERVATION.BE. (2022). Observations.be. Récupérée 10 mars 2022, à partir de <https://observations.be/>
- OISEAUX.NET. (2022). Les oiseaux. Récupérée 27 février 2022, à partir de <https://www.oiseaux.net/>
- OJA, T., ALAMETS, K. & PÄRNAMETS, H. (2005). Modelling bird habitat suitability based on landscape parameters at different scales. *Ecological Indicators*, 5(4), 314-321. <https://doi.org/10.1016/j.ecolind.2005.03.008>
- P. McCULLAGH & J. A. NELDER. (1989). *Generalized linear models. Second edition* (Chapman and hall).
- PANNEKOEK J. & VAN STRIEN ARCO J. (2005). TRIM 3 manual (TRends and Indices for Monitoring data). Récupérée 16 mars 2022, à partir de https://www.researchgate.net/publication/270100638_TRIM_3_manual_TRends_and_Indices_for_Monitoring_data

- PAUL GAILLY. (s. d.). Ethologie - Le calendrier des oiseaux.
- PECBMS. (2022). Récupérée 17 avril 2022, à partir de <https://pecbms.info/>
- PETER A. BURROUGH & A. FRANK. (1996). *Geographic Objects with Indeterminate Boundaries* (1st). Récupérée 29 mai 2022, à partir de <https://www.routledge.com/Geographic-Objects-with-Indeterminate-Boundaries/Burrough-Frank/p/book/9780748403875>
- PlantNET - FloraOnline. (s. d.). Récupérée 20 février 2022, à partir de <https://plantnet.rbgsyd.nsw.gov.au/cgi-bin/NSWfl.pl?page=nswfl&lvl=sp&name=Leionema~ralstonii>
- Portail Biodiversité en Wallonie : espèces, biotopes, sites de grand intérêt biologique, actualités. (s. d.). Récupérée 27 février 2022, à partir de <http://biodiversite.wallonie.be/fr/accueil.html?IDC=6>
- POTTS, J. M. & ELITH, J. (2006). Comparing species abundance models. *Ecological Modelling*, 199(2), 153-163. <https://doi.org/10.1016/j.ecolmodel.2006.05.025>
- PURVIS, A. & HECTOR, A. (2000). Getting the measure of biodiversity. *Nature*, 405(6783), 212-219. <https://doi.org/10.1038/35012221>
- RADOUX, J., BOURDOUXHE, A., COOS, W., DUFRÊNE, M. & DEFOURNY, P. (2019). Improving Ecotope Segmentation by Combining Topographic and Spectral Data [Number : 3 Publisher : Multidisciplinary Digital Publishing Institute]. *Remote Sensing*, 11(3), 354. <https://doi.org/10.3390/rs11030354>
- Réserves Naturelles RNOB / Natagora (organisme) - Réseau IDée asbl. (s. d.). Récupérée 15 novembre 2021, à partir de https://www.reseau-idee.be/adresses-utiles/fiche.php?org_id=1138
- SANZ-PÉREZ, A., SOLLMANN, R., SARDÀ-PALOMERA, F., BOTA, G. & GIRALT, D. (2020). The role of detectability on bird population trend estimates in an open farmland landscape. *Biodiversity and Conservation*, 29(6), 1747-1765. <https://doi.org/10.1007/s10531-020-01948-0>
- SCHÖPFER, E., LANG, S. & ALBRECHT, F. (2008). Object-fate analysis - spatial relationships for the assessment of object transition and correspondence [ISSN : 1863-2246 Series Title : Lecture Notes in Geoinformation and Cartography]. In T. BLASCHKE, S. LANG & G. J. HAY (Éd.), *Object-Based Image Analysis* (p. 785-801). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77058-9_43
- SEGURADO, P. & ARAÚJO, M. B. (2004). An evaluation of methods for modelling species distributions : Methods for modelling species distributions. *Journal of Biogeography*, 31(10), 1555-1568. <https://doi.org/10.1111/j.1365-2699.2004.01076.x>
- STAT 501 - Regression Methods. (s. d.). Récupérée 12 novembre 2021, à partir de <https://online.stat.psu.edu/stat501/lesson/welcome-stat-501>

- STEHMAN, S. V. & WICKHAM, J. D. (2011). Pixels, blocks of pixels, and polygons : Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment*, 115(12), 3044-3055. <https://doi.org/10.1016/j.rse.2011.06.007>
- SUTHERLAND, W. J., NEWTON, I. & GREEN, R. (Éd.). (2004). *Bird ecology and conservation : a handbook of techniques* [OCLC : ocm56013902]. Oxford University Press.
- VER HOEF, J. M. & BOVENG, P. L. (2007). QUASI-POISSON VS. NEGATIVE BINOMIAL REGRESSION : HOW SHOULD WE MODEL OVERDISPERSED COUNT DATA ? *Ecology*, 88(11), 2766-2772. <https://doi.org/10.1890/07-0043.1>
- VIGIE-NATURE. (s. d.). Vigie-Nature. Récupérée 2 juin 2022, à partir de <https://www.vigienature.fr/fr>
- VORISEK, P. (2008). *A best practice guide for wild bird monitoring schemes* [OCLC : 612067964]. Czech Society for Ornithology CSO.
- WALLONIE-BRUXELLES, L. F. (2019). *Rapport d'activité du second semestre 2019* (rapp. tech.).
- WEDDERBURN, R. W. M. (1974). Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss-Newton Method. *Biometrika*, 61(3), 439. <https://doi.org/10.2307/2334725>
- WELSH, A., CUNNINGHAM, R., DONNELLY, C. & LINDENMAYER, D. (1996). Modelling the abundance of rare species : statistical models for counts with extra zeros. *Ecological Modelling*, 88(1-3), 297-308. [https://doi.org/10.1016/0304-3800\(95\)00113-1](https://doi.org/10.1016/0304-3800(95)00113-1)

Cartographie à haute résolution des populations d'oiseaux nicheurs en Brabant wallon (Kaye Marvin)

Dans de nombreux endroits du globe, les habitats naturels de l'avifaune sont détruits suite au développement des activités anthropiques. La Wallonie ne fait pas exception à la règle puisque depuis les années 1990, une perte moyenne de 1 % des effectifs par an est observée. Au vu de ce déclin, la conservation et la planification de la gestion des espèces est essentielle. Pour cela, Aves, le pôle ornithologique de Natagora, contribue au recensement des espèces dans le cadre du programme de Suivi des Oiseaux Communs en Wallonie (SOCWAL). Ce mémoire s'inscrit dès lors dans le projet « Atlas européen », dont Aves fournit les données concernant la Région wallonne avec pour objectif d'obtenir une cartographie à haute résolution de l'abondance des populations d'oiseaux.

Pour mener à bien cette étude, la position géographique de chaque observation est enregistrée lors de la campagne de terrain. Cette position, directement reliée à un écotope, permet d'associer à l'individu les variables quantitatives constituant l'écotope. Ensuite, grâce à ces variables, un modèle prédictif suivant une régression de Poisson calcule les densités de chaque espèce par écotope permettant de cartographier l'abondance de l'espèce. Enfin, une validation des résultats et une analyse de sensibilité sont également réalisées dans le cadre de cette étude.

Ce travail se basant sur une seule campagne de terrain, la difficulté principale réside dans la faible quantité de données récoltées. Ceci appuie tout l'intérêt des plateformes de sciences participatives qui ont permis de récolter un nombre non négligeable d'informations. Au total, plus de 4000 individus, dont 41 espèces différentes, ont été recensés. Sur ces 41 espèces, notre modèle a pu estimer l'abondance par écotope pour 9 d'entre-elles dont 2 sont représentatives de la réalité de terrain. Cependant, l'analyse de sensibilité a permis d'identifier certains critères permettant d'améliorer considérablement les résultats.

Cette étude démontre que renseigner la position exacte de l'observation apporte une réelle plus-value aux résultats. La résolution spatiale des cartographies est passée du kilomètre carré (Aves) à une résolution calibrée sur les écotopes dont la superficie est de l'ordre de l'hectare, soit 100 fois plus petit. Dès lors, ces cartographies plus précises pourraient améliorer les plans de gestion et de conservation de l'avifaune.