



**HAL**  
open science

# Evaluation des tailles de populations des oiseaux communs de France par les données de sciences participatives

Jean Nabias

## ► To cite this version:

Jean Nabias. Evaluation des tailles de populations des oiseaux communs de France par les données de sciences participatives. Biodiversité. Museum national d'histoire naturelle - MNHN PARIS, 2024. Français. ⟨NNT : 2024MNHN0028⟩. ⟨tel-05025280⟩

**HAL Id: tel-05025280**

**<https://theses.hal.science/tel-05025280v1>**

Submitted on 8 Apr 2025

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

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



HAL Authorization



**MUSEUM NATIONAL D'HISTOIRE NATURELLE**  
Ecole Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2024

N°attribué par la bibliothèque  
2024MNHN0028

**THESE**

Pour obtenir le grade de

**DOCTEUR DU MUSEUM NATIONAL D'HISTOIRE NATURELLE**

Spécialité : Science de la Conservation

Présentée et soutenue publiquement par

**Jean Nabias**

**Le 12 novembre 2024**

---

**Evaluation des tailles de populations d'oiseaux  
communs de France par les données de sciences  
participatives**

---

Sous la direction de : **Emmanuelle Porcher (HDR), Luc Barbaro (HDR), Romain  
Lorrillière, Benoit Fontaine, Jérémy Dupuy et Laurent Couzi**

**JURY :**

Mme Céline Clauzel	Professeure, Université Paris 1, Paris (075), France	Présidente du jury
M. Marc Kéry	Directeur de recherche, Swiss Ornithological Institute, Sempach, Suisse	Rapporteur
Mme. Charlotte Francesiaz	Chargée de recherche, Office Français de la Biodiversité (OFB), France	Examinatrice
M. Bruno Faivre	Professeur, Université de Bourgogne, Dijon, France	Examinateur
Mme. Emmanuelle Porcher	Professeure, Museum national d'Histoire naturelle, CESCO, Paris, France	Directrice de thèse
M. Romain Lorrillière	Ingénieur de recherche, CNRS, Paris, France	Encadrant de thèse
M. Luc Barbaro	Chargé de recherche, INRAE, Dynafor, Toulouse, France	Membre invité
M. Benoît Fontaine	Ingénieur de recherche, OFB-PATRINAT, CESCO, Paris, France	Membre invité
M. Jérémy Dupuy	Responsable du projet 'Oiseaux de France', LPO, Rochefort, France	Membre invité
M. Laurent Couzi	Responsable du Service Connaissance, LPO, Rochefort, France	Membre invité

# Résumé

---

L'abondance des individus au sein des populations constitue une information essentielle dans l'évaluation et le suivi de la biodiversité. La fiabilité de son estimation nécessite la prise en compte de l'ensemble des facteurs affectant le nombre d'individus détecté, qu'ils soient liés aux conditions environnementales, soit le processus écologique, ou à l'effort de collecte, soit le processus d'observation. Cette thèse a pour objectif général de produire des estimations fiables des effectifs d'oiseaux nicheurs de France métropolitaine afin d'améliorer la connaissance quantitative de l'avifaune française. Pour ce faire, nous avons utilisé des données de sciences participatives provenant de protocoles et suivis de l'avifaune renseignant les distances d'observations.

Dans un premier temps, nous avons mis en place une démarche de modélisation globale des effectifs nicheurs pour 63 espèces d'oiseaux communs en France métropolitaine à l'aide de données collectées par un protocole structuré, renseignant les distances d'observations des individus détectées selon des répliques temporels. En comparant ces estimations actualisées avec celles du précédent atlas des oiseaux nicheurs, qui ignorait les variations liées à l'effort de collecte, nous avons constaté que 37 espèces étudiées étaient sous-estimées en l'absence de prise en compte du processus d'observation.

Dans la première partie, les données utilisées provenaient d'un jeu de données structuré qui, en raison de l'aspect volontaire des sciences participatives est un compromis entre d'une part la collecte d'observations demandant un effort de collecte important, par exemple, la réalisation de nombreux répliques temporels sur un faible nombre de sites, et d'autre part la collecte d'observations demandant peu d'effort sur une étendue spatiale bien plus grande. Dans un second temps, nous avons donc développé un modèle hiérarchique

## *Résumé*

intégré permettant d'utiliser conjointement deux jeux de données de sciences participatives distincts. Mes travaux ont démontré que l'intégration de ces données, prenant en compte les spécificités de chaque jeu, permettait de réduire les incertitudes associées aux paramètres liés aux deux processus jouant sur les estimations des modèles, le processus d'observation et le processus écologique.

Dans la dernière partie, je montre l'importance de prendre en compte le processus d'observation pour estimer de manière fiable le nombre d'individus pour estimer correctement les métriques de diversité taxonomique, phylogénétique et fonctionnelle des communautés d'oiseaux. Les résultats obtenus semblent confirmer et étendre les conclusions des deux premières parties au niveau des populations, et tendent à montrer que l'omission de la modélisation du processus d'observation peut entraîner des estimations biaisées des indices de diversité pouvant conduire à des représentations incomplètes des communautés, notamment de leurs caractéristiques phylogénétiques et fonctionnelles.

Mots-clés : Suivi des populations ; Taille de population ; Sciences participatives ; Distance sampling

# Abstract

---

Species abundance is an essential variable for assessing and monitoring biodiversity. Reliable estimates require taking account of all factors affecting the number of detected individuals, whether related to environmental conditions, *i.e. the ecological process*, or related to collection effort, *i.e. the observation process*. The main objective of this thesis is to improve knowledge of French avifauna by producing reliable estimates of breeding bird numbers in metropolitan France. To achieve this goal, we relied on citizen science data from bird monitoring schemes that recorded observation distances.

In the first part, we implemented a global modelling framework to estimate the population size of 63 common breeding birds using data provided from a structured scheme that provided observation distances of detected individuals over several temporal replicates. Comparing these newer estimations to those of the previous atlas, which did not account for the observation process, we found that 37 species were underestimated.

The data used in the first part came from a structured scheme that due to the voluntary aspect of citizen science faced a trade-off between (i) collecting data requiring significant effort, such as temporal replicates at a few sites, and (ii) gathering data with less effort but over a wider spatial extent. To address this trade-off, we developed an integrated hierarchical model enabling the joint use of two distinct citizen science schemes. Results showed that data integration led to a reduction of uncertainties of parameters linked to observation and ecological processes.

Finally, I focused my research on community indices and the necessity of considering the observation process to estimate taxonomic, phylogenetic and functional diversity metrics reliably. Preliminary results confirm and extend conclusions from the previous parts, showing

## *Abstract*

that neglecting the observation process can lead to biased diversity estimates and misrepresentation of communities, particularly regarding their phylogenetic and functional facets.

Keywords : Population monitoring ; Population size ; Citizen science ; Distance sampling

# Remerciements

---

Je tenais à remercier **Céline Clauzel, Marc Kéry, Charlotte Francesiaz et Bruno Faivre** d'avoir accepté de faire partie de mon jury de thèse et d'avoir accepté de prendre de leur temps pour évaluer mes travaux.

Merci aussi, aux membres de mon comité de suivi de thèse, **Karine Princé, Jean-Yves Barnagaud et Laurent Godet**. Merci pour leur temps, écoute et conseils.

Un grand merci à ma grande équipe d'encadrants. **Emmanuelle**, sans qui je n'aurais pas pu réaliser cette thèse dans ces conditions. **Luc**, pour les échanges constructifs, ton écoute, tes conseils et ta bienveillance tout au long de cette thèse. **Romain**, ton expertise et nos échanges qu'ils furent techniques ou non ont été des moteurs infailibles. **Benoit** pour ton point de vue ornitho, un bon coup de rappel loin des chiffres, formules, interfaces et claviers. **Jérémy et Laurent**, pour votre expertise et sympathie. Merci à vous quatre d'avoir cru en moi et d'avoir osé l'aventure CIFRE après mon stage de master 2.

Je tenais aussi à remercier **l'ensemble des observateurs** qui ont participé aux protocoles et suivis que j'ai pu utiliser durant ma thèse. Merci pour vos efforts, j'espère vous en rendre une infime partie à travers ces travaux et ce manuscrit.

Un immense merci à l'ensemble des membres du CESCO, leur bonne humeur et leur convivialité. Merci pour ce concentré de joie de vivre permanent. Merci à l'ensemble des habitants du 135 et du 61 str ("**Alexandre, Anne, Anne-Christine, Charles, Colin, Emilienne, Fabien, Fanny, Frédéric, Greg, Jaad, Jeanne, Jérôme, Josquin, Juliette, Laura, Léa M., Lise, Lou, Marie Si, Marie Sp, Marine, Martin, Mathilde, Milena, Nelly, Pablo, Sandra, Simon, Solène, Suzanne, Thibault Ga, Thibault Ge, Tiphaine, Vincent, Yuping, Yves, ...**")

Merci aussi aux **membres du pôle connaissance de la LPO** pour leur accueil chaleureux et cela malgré ma faible présence dans les locaux de Rochefort.

Merci à mes co-bureaux **Camille et Perrine**, pour leur attention et bonne humeur.

Merci à toi **Ahmal**, alias la délégation suisse du bureau.

## *Remerciements*

Merci à toi **Léa A.** pour ta sympathie quotidienne (quand j'étais au bureau) et bon courage pour tes ours pyrénéens.

Merci à toi **Ludo**, pour ta bienveillance (assez singulière), pour tous ces moments de détente et le fait d'être l'éternel second à Mario kart.

Merci à toi **Clément**, cette thèse s'est déroulée à une vitesse éclair, on aura réussi à faire rimer émulation scientifique avec rigolade et bon temps.

Merci à **Karl** et **Supangan**, on aura fait un beau chemin depuis la prépa.

Merci à vous deux, **Jayson** et **Valentin**, pour ces soirées et tous ces moments, ce furent de véritables escapades loin des tracasseries de la thèse

Merci à l'ensemble de ma famille, **mon père, ma mère, Julie** et **Garfield**. Vous avez été un véritable pilier tout au long de ma scolarité. Cette thèse n'aurait pas eu lieu sans vous.  
Merci **maman**, tu auras été ma septième encadrante pour cette thèse.

## Discours du président Drek:

---

*‘Très chers habitants de [Novalis], nous les blargs avons un léger problème.*

*Notre planète est devenue totalement inhabitable.*

*Elle est polluée, surpeuplée et intoxiquée.*

*Mais moi, président Drek, j’ai LA solution.*

*Grâce aux meilleurs composants planétaires disponibles, nous travaillons  
actuellement à un monde parfait.*

*Vous vous demandez certainement pourquoi cela vous concerne ?*

*Grâce à une technologie ultra sophistiquée, que vous ne comprendriez sûrement pas,  
nous allons extraire tout ce que nous pouvons de votre planète, pour l’intégrer  
à la nôtre.*

*Hélas, ce changement massif va entraîner un petit dérèglement de votre  
planète. Elle dérivera jusqu’au soleil, où elle explosera et se transformera en  
une boule de gaz incandescente, mais certains sacrifices sont nécessaires...*

*Merci de votre coopération.’*

*Ratchet & Clank, 2002*

# Glossaire

---

**Population** : Ensemble d'individus d'une même espèce qui coexistent à un moment et un espace donné.

**Occurrence** : Observation d'un individu à un temps et lieu précis.

**Abondance** : Nombre d'individus détectés d'une espèce.

**Densité** : Abondance décrite selon l'aire d'une surface.

**Indices** : Synthétisation d'une ou plusieurs facettes de la biodiversité à l'aide de mesures quantitatives. Ils permettent d'évaluer les variations de la biodiversité face à un ensemble d'acteurs en les comparant à différentes localités et/ou périodes temporelles.

**Échantillonnage** : Collecte d'informations, dans notre cas l'identification et le positionnement des individus détectés durant un temps déterminé. Il permet de collecter des informations sur un sous-ensemble des populations ciblées.

**Processus écologique** : Ensemble des facteurs décrivant la distribution spatiale et/ou temporelle d'une variable d'intérêt, dans notre cas l'abondance d'individus.

**Processus d'observation** : Ensemble de facteurs décrivant les variations liées à la collecte de données. Par exemple, il peut comprendre les variations liées à (i) la répartition temporelle de l'échantillonnage, (ii) l'obstruction visuelle et/ou sonore du milieu, (iii) l'hétérogénéité de l'expertise des observateurs.

**Disponibilité** : Probabilité qu'un individu soit à la fois présent dans la zone prospectée par l'observateur, lors de la collecte, mais aussi qu'il manifeste un comportement le rendant détectable, par exemple par l'émission d'un chant.

**Délectabilité** : Probabilité qu'un individu disponible à la détection soit perçu par l'observateur lors de la collecte.

**Modèles hiérarchiques** : Modèles statistiques où la distribution d'un, voire d'une partie des paramètres estimés, est conditionnée par la distribution d'un autre ensemble de paramètres. Ils s'apparentent à une séquence de modèles simples, utilisés pour représenter un système complexe en le subdivisant en sous-systèmes distincts (Kéry & Royle, 2015).

# Table des figures

---

- [Figure 1](#) : Illustration d'un couple de Tourte voyageuse (*Ectopistes migratorius*)
- [Figure 2](#) : Illustration d'une session de chasse de tourte voyageuse
- [Figure 3](#) : Illustration d'autres espèces d'oiseaux nord-américaines éteintes à la fin du 19e siècle
- [Figure 4](#) : Extinctions d'origines humaines récentes. Copyright *American Association for the Advancement of Science*, diffusion autorisée pour ce document unique
- [Figure 5](#) : L'interaction des activités humaines et les changements globaux induits
- [Figure 6](#) : Interaction des différents acteurs gouvernant la mise en place de nouvelles zones protégées
- [Figure 7](#) : Construction et objectif des EBVs
- [Figure 8](#) : L'indice planète vivante (LPI) global entre 1970 et 2018
- [Figure 9](#) : Représentation d'une tranche de la Sierra Nevada, Grinnel 1924
- [Figure 10](#) : Tableau dénombrant le nombre de captures de tourte voyageuse en 1879
- [Figure 11](#) : Apport des différentes sources de données pour l'étude de Galliformes
- [Figure 12](#) : Illustration des facteurs limitant l'efficacité et la précision des méthodes de décompte. Copyright *Oxford University Press*, diffusion autorisée pour ce document
- [Figure 13](#) : Représentation de différents types de relations entre un indice relatif et son paramètre d'intérêt. Copyright *Elsevier*, diffusion autorisée pour ce document
- [Figure 14](#) : Représentation des processus stochastiques sous-jacents à l'estimation d'abondance
- [Figure 15](#) : Visualisation des différentes probabilités du processus d'observation selon la quantité d'individus détectée
- [Figure 16](#) : Ensemble d'études montrant les relations entre l'abondance des oiseaux et les services écosystémiques
- [Figure 17](#) : Représentation schématique de la réalisation d'un point d'écoute STOC. Diffusion autorisée par Axelle Grégoire
- [Figure 18](#) : Représentation schématique des trois jeux de données.
- [Figure 19](#) : Représentation d'un modèle spatial d'état selon la position des individus. Copyright *Springer London*, diffusion autorisée pour ce document
- [Figure 20](#) : Représentation schématique de différents types d'échantillonnage. Copyright *John Wiley & Sons*, diffusion autorisée pour ce document
- [Figure 21](#) : Représentation de la fonction de détection
- [Figure 22](#) : Figures représentant le contexte géographique et la diversité climatique et de types d'habitats en France métropolitaine.
- [Figure 23](#) : Représentation schématique d'un espace environnemental
- [Figure 24](#) : Représentation schématique de différents cas d'étude pour l'inférence d'un paramètre
- [Figure 25](#) : Représentation schématique de la théorie de l'information. Copyright *John Wiley & Sons*, diffusion autorisée pour ce document
- [Figure 26](#) : Distribution spatiale des localités (listes et sites) des dispositifs EPOC et EPOC-ODF entre 2021 et 2023.
- [Figure 27](#) : Ensemble des éléments affectant la probabilité de détection des espèces. Copyright *Elsevier*, diffusion autorisée pour ce document
- [Figure 28](#) : Exemple du cycle de vie d'un passereau. Copyright *John Wiley & Sons*, diffusion autorisée pour ce document
- [Figure 29](#) : Graphique d'un exemple de modèle de population intégré. Copyright *John Wiley & Sons*, diffusion autorisée pour ce document
- [Figure 30](#) : Proposition d'amélioration des critères de menace des listes rouges de l'IUCN
- [Figure 31](#) : Distribution d'un ensemble d'indicateurs selon leurs dimensions scientifique
- [Figure 32](#) : Représentation du LPI selon différentes échelles fonctionnelles et spatiales
- [Figure 33](#) : Fiabilité des données utilisées pour calculer le LPI en fonction des régions de l'IPBES
- [Figure 34](#) : Distribution de l'effet de la croissance des oisillons selon la taille de la couvée étudié par différentes équipes de recherche

# Table des matières

---

<b>Résumé .....</b>	<b>1</b>
<b>Abstract .....</b>	<b>3</b>
<b>Remerciements.....</b>	<b>5</b>
<b>Glossaire .....</b>	<b>8</b>
<b>Table des figures .....</b>	<b>10</b>
<b>Table des matières .....</b>	<b>11</b>
<b><u>I</u>    <b>Introduction générale.....</b></b>	<b>15</b>
I.A    La fragilité du commun : Insights from oblivion.....	15
I.B    Vivant & Changements globaux.....	19
<i>I.B.i</i> <i>Crise de la biodiversité.....</i>	<i>19</i>
<i>I.B.ii</i> <i>Les indices : outils de dialogue, communication et diffusion .....</i>	<i>21</i>
I.C    Les espèces et leurs interactions : comment les caractériser ?.....	26
I.D    L’observation : un filtre nécessaire à l’étude du vivant.....	31
<i>I.D.i</i> <i>Transformation des registres et pratiques de collecte.....</i>	<i>31</i>
<i>I.D.ii</i> <i>Le suivi des effectifs d’oiseaux nicheurs.....</i>	<i>35</i>
<i>I.D.iii</i> <i>Le vivant caché par l’observation .....</i>	<i>39</i>
I.E    Intérêt de conservation : la place du commun .....	45
I.F    Objectifs et questions de recherches .....	49
<b><u>II</u>    <b>Jeux de données et Méthodes.....</b></b>	<b>53</b>
II.A    Jeux de données .....	53
<i>II.A.i</i> <i>STOC-EPS (Suivi Temporel des Oiseaux Communs par</i> <i>Échantillonnage Ponctuel Simple).....</i>	<i>53</i>
<i>II.A.ii</i> <i>EPOC (Estimation des Populations d’Oiseaux Communs) .....</i>	<i>56</i>
<i>II.A.iii</i> <i>EPOC-ODF : Un protocole structuré pour l’estimation des</i> <i>populations d’oiseaux communs .....</i>	<i>57</i>
II.B    Méthodes d’analyses.....	60
<i>II.B.i</i> <i>Hierarchical Distance Sampling : L’estimation d’abondance à partir</i> <i>de données de distance.....</i>	<i>60</i>
<i>Boîte 1: Approches d’estimations basées sur l’échantillonnage et les modèles –</i> <i>Design-based et Model-based.....</i>	<i>66</i>

## Table des matières

<i>Boxe 2: Distance Sampling Hiérarchique prenant en compte l'émigration temporaire</i> .....	67
II.B.ii <i>Environnement et modèles d'études</i> .....	68
II.B.ii.a <i>Espèces cibles</i> .....	68
II.B.ii.b <i>Spécificité du territoire français métropolitain</i> .....	69
II.B.ii.c <i>Réduction des données environnementales et quantification de l'extrapolation</i> .....	71
II.B.iii <i>Approche d'estimation des paramètres</i> .....	74
II.B.iv <i>Diagnostic des modèles</i> .....	77
II.B.v <i>Sélection des modèles</i> .....	78
<b>III <u>Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives</u></b> .....	<b>82</b>
III.A Contexte .....	82
III.B Article .....	84
<i>Abstract</i> .....	85
<i>Introduction</i> .....	86
<i>Materials &amp; Methods</i> .....	90
<i>Results</i> .....	102
<i>Discussion</i> .....	106
III.C Supplementary information .....	116
III.D References.....	143
<b>IV <u>Chapitre 2 : Développement de modèle intégré dans le but d'améliorer le suivi de l'avifaune nicheuse</u></b> .....	<b>154</b>
IV.A Contexte .....	154
IV.B Article .....	157
<i>Abstract</i> .....	158
<i>Introduction</i> .....	158
<i>Material and Methods</i> .....	163
<i>Results</i> .....	171
<i>Discussion</i> .....	176
IV.C Supplementary information .....	181
IV.D References.....	191
<b>V <u>Chapitre 3 : Effet de la non prise en compte du processus d'observation dans l'estimation d'indice de biodiversité</u></b> .....	<b>197</b>
V.A Contexte .....	197

## Table des matières

V.B	Article .....	201
	<i>Abstract</i> .....	202
	<i>Introduction</i> .....	203
	<i>Material and methods</i> .....	206
	<i>Results</i> .....	216
	<i>Discussion</i> .....	220
V.C	Supplementary information .....	225
V.D	References.....	233
<b><u>VI</u></b>	<b><u>Discussion générale .....</u></b>	<b><u>243</u></b>
VI.A	Rappel des résultats.....	243
VI.B.....	Estimation d'abondance : vers une vision plus détaillée de l'avifaune française .....	245
VI.C	Indices et narratifs : une association nécessaire.....	248
	<i>VI.C.i Indicateurs : un ensemble de critères et de tensions</i> .....	248
	<i>VI.C.ii L'agrégation d'indices multi-espèces : cas du LPI</i> .....	251
VI.D	Transmission de l'incertitude scientifique face à la prise d'action.....	257
<b><u>VII</u></b>	<b><u>Références de l'introduction générale.....</u></b>	<b><u>261</u></b>
<b><u>VIII</u></b>	<b><u>Références de la discussion générale.....</u></b>	<b><u>281</u></b>
<b><u>IX</u></b>	<b><u>Annexe: Tableau des tailles de populations d'oiseaux communs nicheurs de France métropolitaine.....</u></b>	<b><u>289</u></b>



# I Introduction générale

---

## I.A La fragilité du commun : Insights from oblivion



*‘The Pigeons were still passing in undiminished numbers, and continued to do so for three days in succession. The people were all in arms.’, John James Audubon, 1831*

*‘We place this bird in the totally-extinct class, not only because it is extinct in a wild state, but only one solitary individual, a twenty-year-old female in the Cincinnati Zoological Gardens, now remains alive’, Hornaday, 1913*

Figure 1 : Illustration d’un couple de Tourtes voyageuses (*Ectopistes migratorius* ; mâle en haut et femelle en bas). Réalisée par Louis Agassiz Fuertes et tirée de Mershon 1907

‘Semblable à une éclipse, la lumière du zénith fut obscurcie. L’air était littéralement empli de pigeons’. Telle était la description de la migration de la Tourte voyageuse, ou pigeon migrateur (*Ectopistes migratorius*; voir Figure 1) par John James Audubon au milieu du 19<sup>e</sup> siècle dans l’Ohio (Audubon, 1831). La Tourte voyageuse était une espèce emblématique de l’Amérique du Nord, mentionnée pour la première fois au courant du 16<sup>e</sup> siècle par le navigateur Cartier, où il en dénombrait une infinité (Schorger, 1955). Un siècle plus tard, lors de son voyage, John Josselyn recensait un déclin de l’espèce et ciblait une cause, celle de la chasse (Josselyn, 1674).

Cette espèce fut convoitée et chassée de manière industrielle durant le 19e siècle (Halliday, 1980). Les méthodes de collecte d'individus étaient diverses, à savoir, l'intoxication par

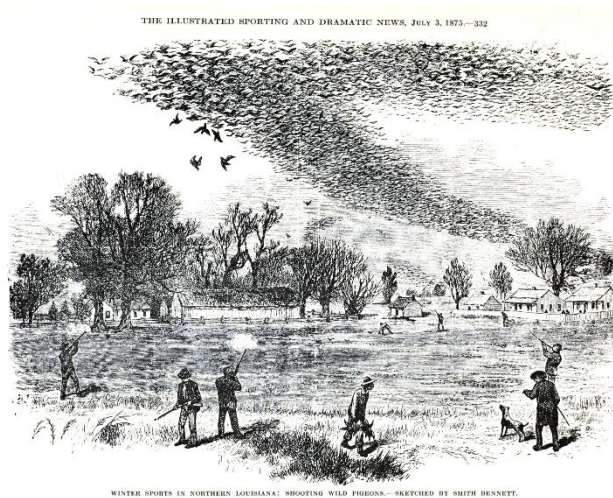


Figure 2 : Illustration d'une session de chasse de tourte voyageuse, durant l'hiver au nord de la Louisiane. Illustration de Smith Bennet dans 'Illustrated Sporting and Dramatic News', 3 Juillet 1875.

combustion de soufre, l'intoxication de graines avec de l'alcool, mais les principaux moyens de collecte étaient la mise en place de filets de capture et le tir (Figure 2; Schorger, 1955).

Cette espèce représentait une ressource nutritive importante pour les amérindiens et colons européens (Stanton, 2014). Bien que les premiers l'aient exploitée à une échelle locale, celle-ci constituait un secteur économique pour les derniers. Le déclin des populations a été accéléré par l'expansion vers l'ouest ainsi que le développement des systèmes ferroviaires et de communications (Halliday, 1980; Stanton, 2014). Son extinction fut provoquée par une chasse non régulée, conduisant à la collecte sans égard d'individus adultes, juvéniles et d'œufs, mais aussi par la perte d'habitat provoquée par la déforestation pour l'expansion de terres agricoles (Blockstein & Tordoff, 1985).

La surexploitation de cette espèce peut avoir pour origine l'aspect providentiel de celle-ci (Audubon, 1831). Malgré l'importante quantité d'individus, on remarque le manque d'estimations fiables, allant d'estimations basées sur des ressentis (Josselyn, 1674) à celles basées sur l'étude de flux d'individus par surface variant de plus de 1 milliard (Audubon,

1831) à plus de deux milliards d'individus (Wilson, 1829). Mais les récits du 17<sup>e</sup> siècle de John Josselyn semblent indiquer plutôt une grande variabilité démographique non seulement liée à la collecte d'individus, mais potentiellement aussi aux fluctuations de la ressource

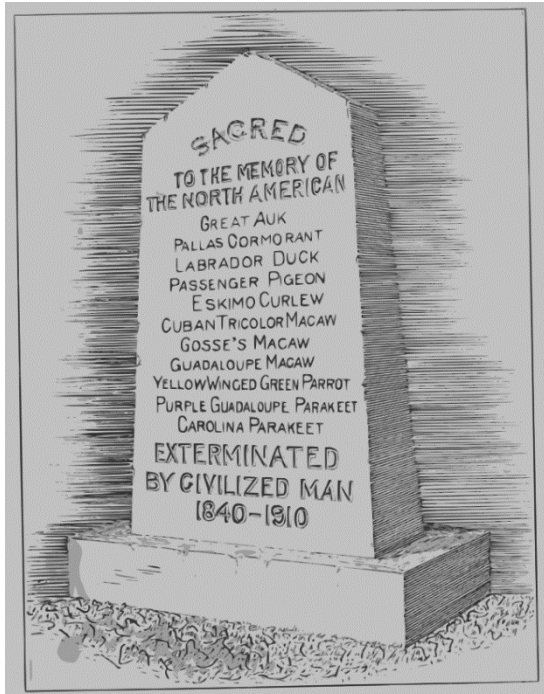


Figure 3 : Illustration d'autres espèces d'oiseaux nord-américaines éteintes à la fin du 19<sup>e</sup> siècle. De haut en bas, Great Auk (Grand pingouin - *Pinguinus impennis*) ; Pallas Cormorant (Cormoran de Pallas - *Urile perspicillatus*) ; Labrador Duck (Eider du Labrador - *Camptorhynchus labradorius*) ; Tourte voyageuse ; Eskimo Curlew (Courlis esquimau - *Numenius borealis*) ; Cuban tricolored Macaw (*Ara tricolor*) ; Jamaican Red Macaw (*Ara gossei*)\* ; Guadeloupe Macaw (*Ara guadeloupensis*)\* ; Yellow-winged green Parrot (*Amazona olivacea*) ; Purple Guadeloupe Parakeet (*Anodorhynchus purpurascens*) ; Carolina Parakeet (Conure à tête jaune - *Conuropsis carolinensis*). Illustration tirée de Hornaday 1913. (\*) : Espèces hypothétiques

alimentaire. L'espèce était connue comme étant une espèce spécialiste granivore fortement dépendante des variabilités spatio-temporelles des productions de fruits d'arbres tels que le Hêtre américain (*Fagus grandifolia*), les chênes (*Quercus* spp.) ou le Châtaignier d'Amérique (*Castanea dentata*; Bucher, 1992).

La disparition de cette espèce, en plus d'autres (Figure 3 ; Scott, Burton & Fitter, 1987) a promu l'établissement de jalons afin de déterminer le statut de conservation des espèces (Fitter & Fitter, 1987). Rétrospectivement, les statuts de conservation actuels (Critères des listes rouges ; IUCN, 2012) auraient été suffisants pour alerter sur le déclin de la Tourte voyageuse (Stanton, 2014). Une représentation de la dissonance entre l'état 'réel' de la population et la représentation de celle-ci peut être décrite par le commentaire du comité de l'État d'Ohio sur la mise en place d'une loi de conservation de la Tourte voyageuse en 1857 :

*‘La Tourte voyageuse ne nécessite d’aucune protection. Extraordinairement prolifique [...] il se trouve ici, aujourd’hui, et sera ailleurs demain, aucun acte de destruction ordinaire ne pourra provoquer son déclin’*(King, 1978). Cette dissonance de représentation entre l’état actuel de la nature et la perception de celle-ci, ou amnésie écologique – *shifting baseline syndrom* (Pauly, 1995) – tend à percevoir la nature comme stable, malgré un déclin, ou un changement structurel constant (Soga & Gaston, 2018). Ce phénomène est accentué par trois facteurs: (i) le manque d’évidence et de données fiables attestant de l’état des populations ; (ii) la perte d’interactions directe avec l’environnement, ou l’extinction d’expérience (Miller, 2005) ; et (iii) la perte de familiarité envers celui-ci, *i.e.* perte de connaissance sur l’histoire naturelle (Pilgrim et al., 2008).

## I.B Vivant & Changements globaux

### I.B.i Crise de la biodiversité

La Tourte voyageuse ne représente qu'un exemple d'espèce parmi beaucoup d'autres et illustre un problème plus large : celui de la crise de la biodiversité et l'entrée dans une nouvelle ère, l'anthropocène (Crutzen & Stoermer, 2000). Ce changement d'ère géologique est marqué par l'intensification des effets des activités humaines sur l'ensemble du vivant et l'altération des flux d'énergies de la planète (Steffen et al., 2011).

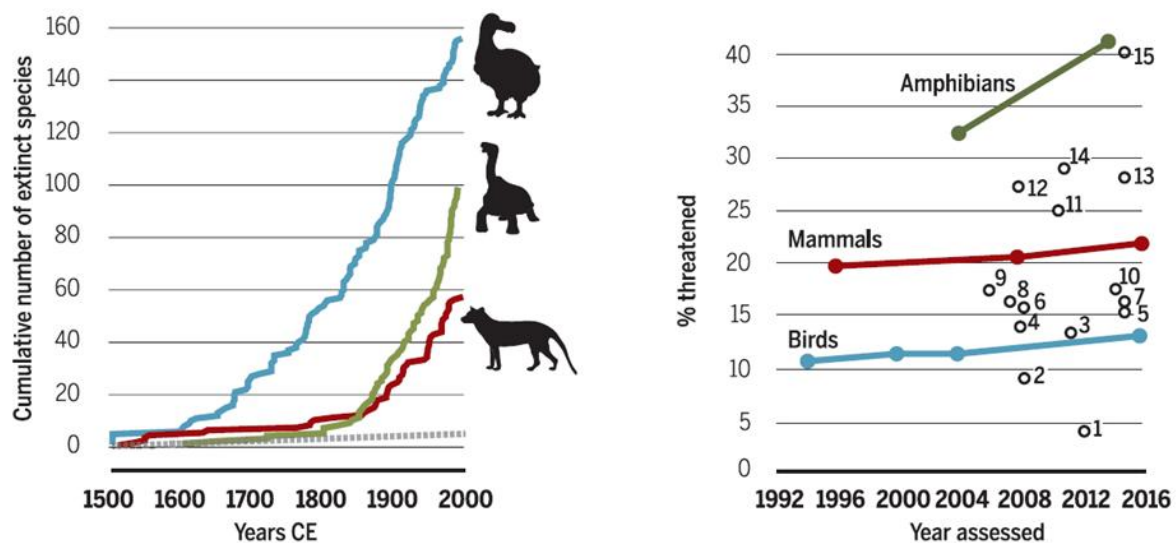


Figure 4 : Extinctions d'origines humaines récentes. A gauche, le nombre cumulé d'espèces de vertébrés éteintes depuis 1500 (UICN – Liste rouge); en bleu les oiseaux; en rouge les mammifères et en vert les autres vertébrés. La ligne pointillée grise correspond au taux d'extinction basal estimé à partir des données fossiles, à savoir deux extinctions par million d'espèces par an (Ceballos et al., 2015). A droite, les tendances récentes de la part des espèces menacées d'extinction pour les oiseaux (bleu), mammifères (rouge) et amphibiens (vert) comparées à la part d'espèces menacées d'autres groupes taxonomiques. (1) Concombres de mer ; (2) Odonates ; (3) Reptiles ; (4) Algues ; (5) Bryophytes ; (6) Mangroves ; (7) Ptéridophytes ; (8) Crabes ; (9) Requins et raies ; (10) Crevettes ; (11) Écrevisses ; (12) Coraux ; (13) Cactus ; (14) Poissons d'eaux douces ; (15) Gymnospermes. Tiré de Johnson et al., 2017.

La crise de la biodiversité (Figure 4) est due à des effets directs et indirects des activités humaines (Figure 5 ; Mace, 2011), à savoir le changement d'usage des sols, au travers de la déforestation, de l'agriculture et de la fragmentation d'habitats (Travis, 2003;

Hoekstra et al., 2005; Newbold et al., 2015), l'introduction d'espèces invasives (Gurevitch & Padilla, 2004), la surexploitation de ressources (Rosser & Mainka, 2002; Pauly, 2008), la pollution (McNeely, 1992), ainsi que les changements climatiques (Travis, 2003; Soroye, Newbold & Kerr, 2020).

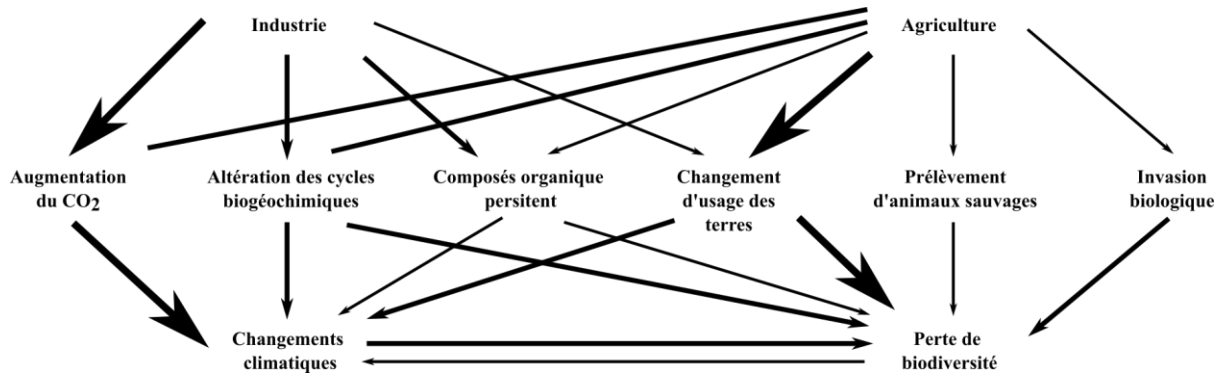


Figure 5 : L'interaction des activités humaines et les changements globaux induits, adapté de Zavaleta & Heller, 2012

La modification des composantes des écosystèmes, par l'altération du milieu physique, ou le retrait -voire l'ajout- d'espèces tend à modifier localement l'abondance et la richesse spécifique des milieux (Hooper et al., 2012; Newbold et al., 2015) et modifie les interactions d'espèces (Tylianakis et al., 2008). Les oiseaux représentent des objets d'études et indicateurs pertinents pour étudier les effets des changements globaux. En effet, ces espèces possèdent des cycles de vies complexes composés de nombreuses interactions diverses au cours de leur existence (Moolna et al., 2018). Plus spécifiquement, cela se traduit par la tendance à favoriser des individus ayant une phénologie de reproduction précoce (Gaüzère & Devictor, 2021) ou un départ en migration précoce (Lehikoinen et al., 2019) possiblement afin de se synchroniser avec la ressource alimentaire (Visser, Holleman & Gienapp, 2006; Parmesan, 2006). Cela peut aussi consister en une modification, réduction ou allongement des voies migratoires de certaines espèces (Berthold, 1999; Marion, Vessem & Ulenaers, 2000; Howes, Symes & Byholm, 2019).

Les effets des changements globaux (Figure 5) se manifestent de manière hétérogène à travers les diverses échelles du vivant—organismes, espèces, populations, communautés, et écosystèmes ; (Dupuis-Tate & Fischesser, 2017), notamment l’interaction entre les conditions climatiques et les habitats provoquant un cocktail anthropogénique (Travis, 2003). Les différents niveaux du vivant possèdent des dynamiques spatio-temporelles différentes (Levin, 1992; Wu, 2013) qui complexifient l’analyse des changements globaux (Bowler et al., 2020). Pour y faire face, la simplification par des indices est souvent nécessaire (voir : *Les indices : outils de dialogue, communication et diffusion*). Ces indices ont pour objectif de synthétiser les éléments<sup>1</sup>, ou produits<sup>2</sup> des écosystèmes selon des proxys de pressions d’activités humaines.

#### I.B.ii Les indices : outils de dialogue, communication et diffusion

Face à la dégradation du vivant provoquée par les activités humaines, des mouvements de conservation de la nature sont apparus à la fin du 18e siècle (Ladle, Jepson & Gilson, 2011). Ces mouvements avaient des origines et objectifs hétérogènes allant de la préservation d’une nature à des fins contemplatives ou pour pérenniser son exploitation (Voir Tableau 2.1 de Ladle, Jepson & Gilson, 2011). Ces mesures sont appliquées afin de protéger la nature (Figure 6) pour soit (i) sa valeur intrinsèque, considérant que les non-humains possèdent les mêmes droits que les humains à prospérer de façon pérenne ; (ii) sa valeur instrumentale, prenant en compte les bénéfices que sa préservation pourrait apporter à l’Homme, principalement financiers, mais aussi sociaux – *les services écosystémiques* (Hornaday, 1913; Ehrlich & Mooney, 1983; Ehrlich & Ehrlich, 1992) ; ou (iii) sa valeur relationnelle, considérant les interactions non transposables liant les humains à la nature, *i.e.*

---

<sup>1</sup> Par exemple, les espèces, populations ou communautés, mais aussi les habitats.

<sup>2</sup> On pourra prendre l’exemple de la productivité primaire suivie par le Normalized Difference Vegetation Index (NDVI).

l'attachement à un lieu provoqué par un souvenir (Himes & Muraca, 2018; Deplazes-Zemp, 2024).

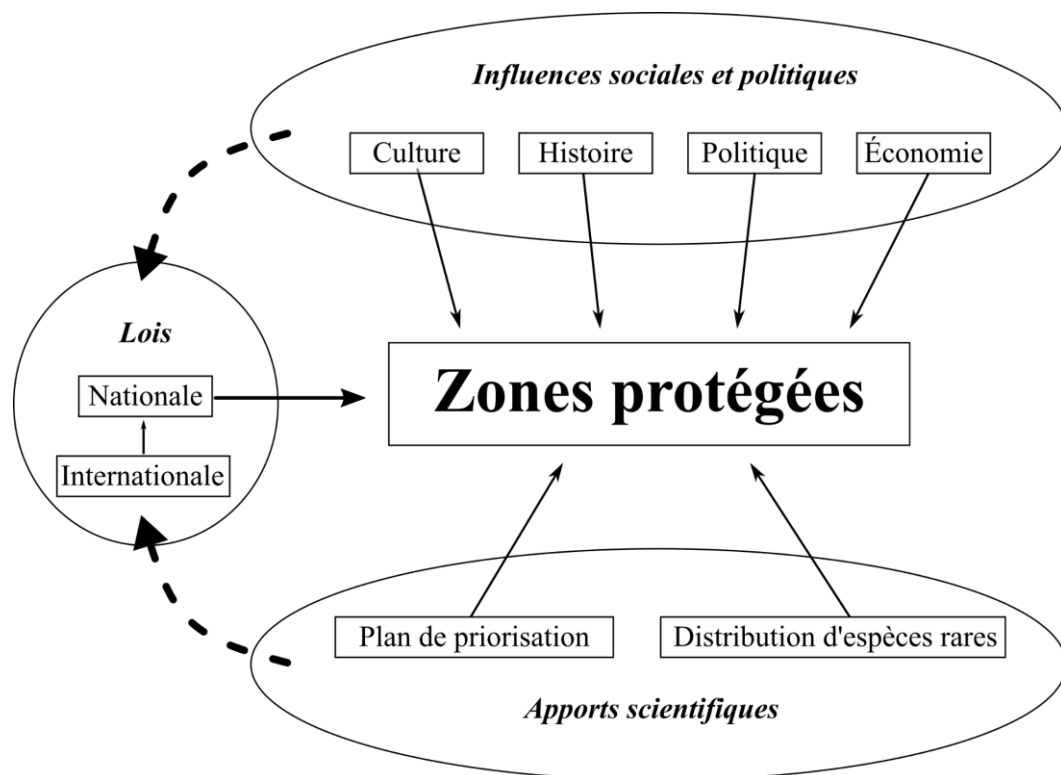


Figure 6 : Interaction des différents acteurs gouvernant la mise en place de nouvelles zones protégées. Adapté de Ladle, Jepson & Gilson, 2011

Le sommet 'Earth summit' des Nations Unies de 1992 à Rio de Janeiro marque un tournant sans précédent sur la prise en compte de la nécessité de protéger la biodiversité face aux impacts des activités humaines (United Nations, 1993). Afin de suivre les engagements des parties prenantes, plusieurs indicateurs, ou métriques ont été développés – les variables essentielles de la biodiversité (EBV ; Pereira et al., 2013), permettant de quantifier, évaluer et gérer les variations des différentes facettes de la biodiversité face aux changements globaux et l'application des mesures de conservation (Wilhere, 2002). Ces EBVs permettent d'assurer le suivi de plusieurs formes de la biodiversité, allant de la composition génétique des populations, de leurs distributions spatiales et d'abondance jusqu'aux fonctions des écosystèmes (Figure 7; Pereira et al., 2013; Jetz et al., 2019). Au vu de la complexité de la

biodiversité, ces indicateurs permettent de synthétiser l'information avec pour principaux objectifs l'étude et la communication des résultats. On pourra citer l'exemple de l'indice planète vivante – *Living Planet Index* (LPI ; Collen et al., 2009)<sup>3</sup> qui correspond à l'agrégation des variations d'abondance de dizaines de milliers de populations de plusieurs milliers d'espèces (Wild Fund for Nature; WWF, 2020) entre 1970-2018 (Figure 8; Almond et al., 2022). Il est utilisé par la Convention sur la Diversité Biologique (CDB) des Nations Unies pour communiquer en objectivant la crise de la biodiversité auprès des décideurs, mais aussi pour le grand public, au travers des rapports de l'organisation non gouvernementale WWF (Almond et al., 2022).

---

<sup>3</sup> La controverse sur son manque de sensibilité et sont une interprétation confuse sera abordé en discussion (voir : *L'agrégation d'indices multi-espèces : cas du LPI*)

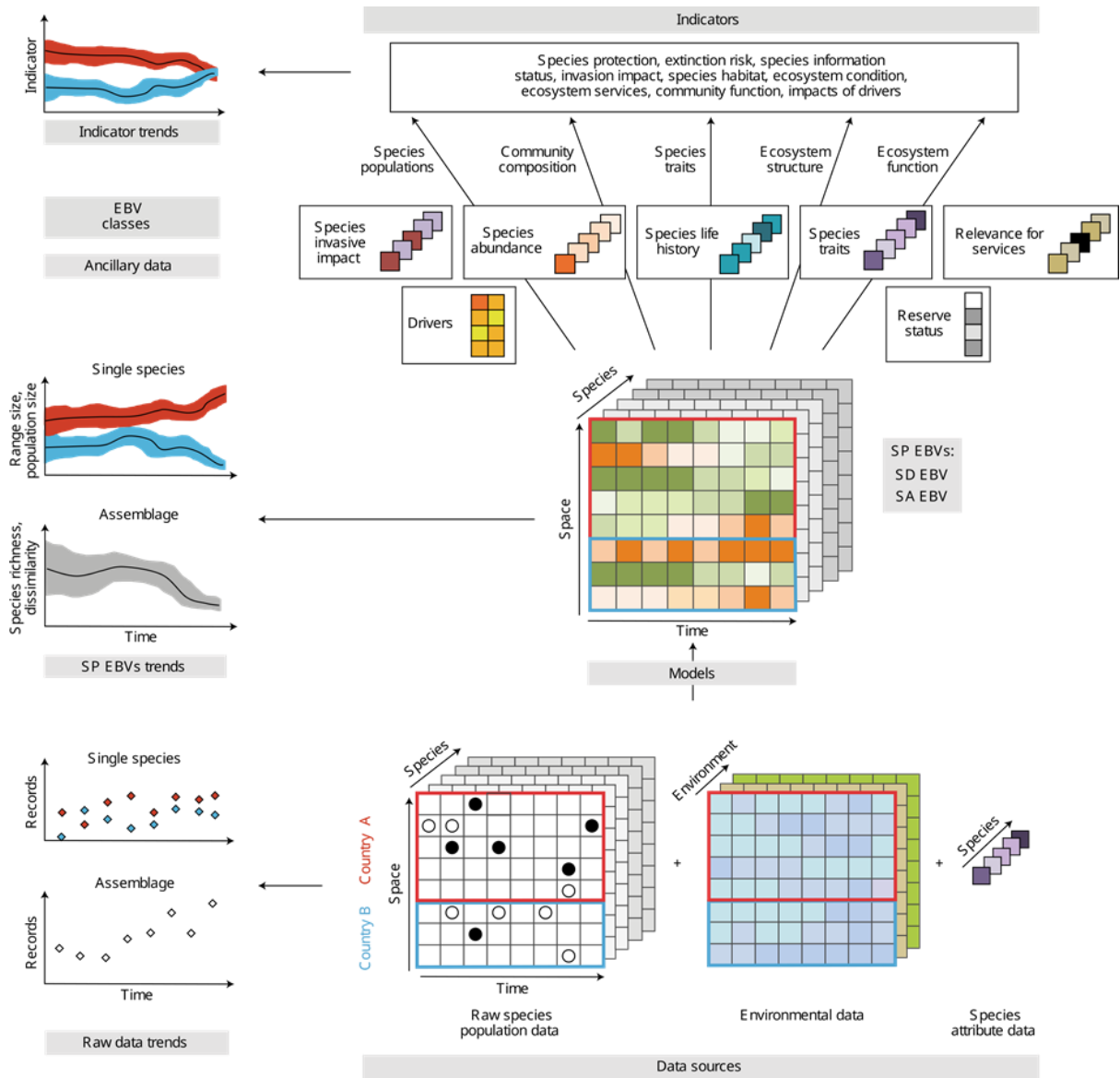


Figure 7 : Construction et objectifs des EBVs. La première étape (bas de figure) dans la construction des EBVs correspond à l’acquisition de données d’occurrence selon des gradients spatiaux et temporels. En bas de figure, la première étape de la construction des EBVs (ou Species Populations EBV – SP EBV) correspond à l’acquisition de données d’occurrence et environnementales sur un ensemble d’espèces, renseignant sur la composition des communautés spatialement et temporellement. Au milieu, ces données sont ensuite utilisées pour calibrer des modèles décrivant ces variations. Les prédictions peuvent varier selon les EBVs et peuvent décrire les distributions d’espèces (Species Distribution EBV – SD EBVs), la distribution spatiale des abondances (Species Abundance EBV – SA EBV). Il est possible d’ajouter des données supplémentaires ‘Ancillary data’ comme des données morphologiques et fonctionnelles de chaque espèce afin d’obtenir une vision plus holiste des communautés. A gauche: il est possible d’agréger les EBVs selon différentes échelles spatiales afin de renseigner sur l’état des populations globalement ou selon certaines zones géographiques, typiquement différents pays ou continents. Figure tirée de Jetz et al., 2019.

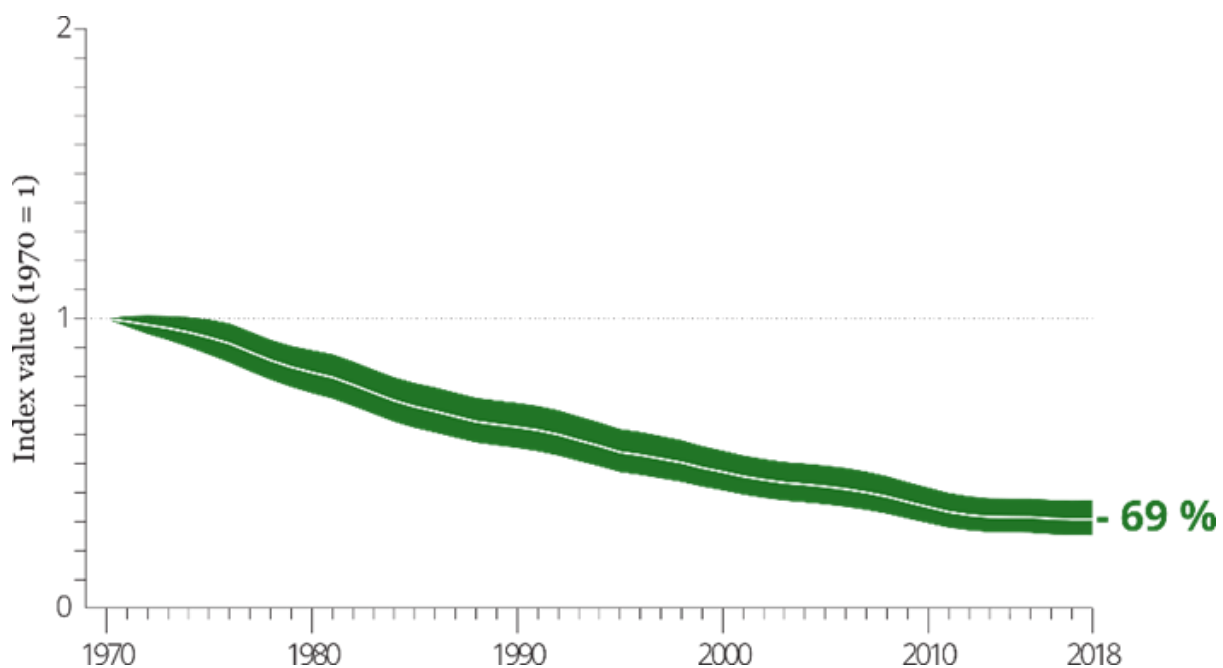


Figure 8 : L'indice planète vivante global (LPI) entre 1970 et 2018. Il représente la variation moyenne de l'abondance relative de 31,821 populations de 5,230 espèces. La courbe blanche représente la tendance moyenne avec son incertitude à 95% en vert. Tiré du rapport 2022 de la WWF (Almond et al., 2022)

I.C Les espèces et leurs interactions : comment les caractériser ?

Dans une optique de conservation, où l'on veut précisément agir sur les facteurs biotiques ou abiotiques permettant de favoriser la pérennité d'une espèce, il est important de déterminer l'ensemble de ces conditions où l'espèce peut persister – *sa niche écologique*. Le concept de niche écologique est multiple. Ainsi, Charles Elton (1900-1991) considère que la niche écologique est en grande partie liée à la taille et aux habitudes alimentaires des espèces. Cette représentation considère que l'espace disponible est principalement régi par des interactions biotiques, typiquement des interactions trophiques – ‘... the "niche" of an animal means its place in the biotic environment, its relations to food and enemies.’ (Elton, 1927). Joseph Grinnel (1877-1939) quant à lui considère que les conditions abiotiques du milieu sont les principaux facteurs déterminant la niche écologique des espèces (Figure 9; Grinnel & Storer, 1924).

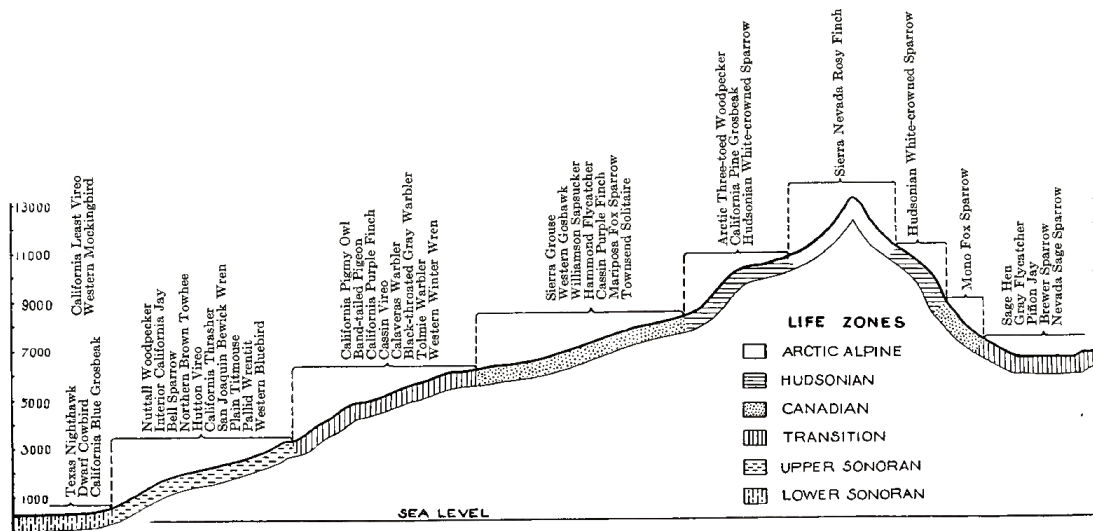


Figure 9 : Représentation d'une tranche de la Sierra Nevada définie en 6 sous-sections – ceintures – regroupant des caractéristiques environnementales spécifiques. Ces zones abritent différents assemblages d'espèces. Tiré de Grinnel 1924.

Bien que Grinnel n'exclut pas le rôle des interactions biotiques, notamment la compétition pour la ressource, dans l'établissement des niches écologiques (Grinnell & Storer, 1924), ce sera Evelyn Hutchinson (1903-1991) qui réalisera une description mathématique alliant les différents concepts de niche écologique (Colwell & Rangel, 2009). D'après Hutchinson, on peut décrire la niche écologique d'une espèce par un hypervolume représentant les étendues des  $N$  conditions environnementales où une espèce peut subsister<sup>4</sup>. De cette représentation, découlent deux niches différentes, (i) la niche fondamentale qui décrit uniquement les interactions abiotiques de l'espèce et (ii) la niche réalisée qui correspond à la surface de l'hypervolume de l'espèce qui n'intersecte pas avec celui des autres espèces (Hutchinson, 1957).

C'est cette représentation de la niche écologique qui sera adoptée en grande majorité en écologie (Slack & Wilson, 2011). Plus particulièrement dans le domaine de la conservation, où la modélisation des habitats favorables aux espèces, en plus des conditions climatiques (Barnagaud et al., 2012), est essentielle pour mettre en place des mesures de conservation adéquates ou évaluer l'état des populations (Sánchez-Cordero et al., 2005; Schwartz, 2012; Santangeli et al., 2022).

Plusieurs approches ont été développées pour quantifier et analyser ces interactions. Certaines méthodes comme l'analyse des réseaux d'interactions trophiques – *Trophic Network Analysis* (Buzhdygan, Patten & Rudenko, 2012) se concentrent sur l'ensemble biotique des écosystèmes, tandis que d'autres approches analysent les flux d'énergie entre les différents éléments des écosystèmes – *Network Analysis* (Fath & Patten, 1999). Dans nos recherches, nous nous sommes restreints aux interactions entre les espèces et leur environnement physique – *l'autecology*. Pour étudier ces interactions, nous avons recouru à

---

<sup>4</sup> Plus précisément, Hutchinson décrit qu'au sein de l'hypervolume, le taux de croissance des espèces est supérieur à 1, ce qui conduit à une persistance ou croissance des populations.

des modèles de distributions d'espèces – *Species Distribution Models* (SDM ; [Elith & Leathwick, 2009](#)), décrivant la variation spatiale d'occurrence ou d'abondance selon des relations linéaires ou non linéaires<sup>5</sup> à un ensemble d'habitats, à l'aide de données d'observations collectées sur le terrain. Lorsqu'une espèce, population ou communauté est suivie sur de longues périodes temporelles, on peut aussi modéliser les variations temporelles du nombre d'individus ([Yoccoz, Nichols & Boulinier, 2001](#); [Magurran et al., 2010](#)) selon différents types de regroupement taxonomiques, fonctionnels ou spatiaux ([Loh et al., 2005](#)).

Récemment, de nouvelles approches comme les modèles de distribution d'espèces jointes – *Joint Species Distribution Models* (JDSM ; [Pollock et al., 2014](#); [Warton et al., 2015](#)) considèrent prendre en compte à la fois les interactions abiotiques et biotiques ([Poggiato et al., 2023](#)). Le sous-jacent de cette approche ([Warton et al., 2015](#)) réside dans (i) l'explication des interactions abiotiques par les relations à l'habitat qui soient linéaires, voire non linéaires et (ii) l'analyse des résidus pour chaque espèce censée détecter des interactions biotiques (mutualisme et compétition pour des corrélations, respectivement, positives et négatives). Cependant, ces corrélations ne correspondent pas forcément à de véritables interactions biotiques ([Blanchet, Cazelles & Gravel, 2020](#)) et peuvent par exemple correspondre à des données environnementales manquantes ([Vallé et al., 2024](#)).

En dehors du cas spécifique des JSDMs, les SDM sont majoritairement employés pour estimer les répartitions d'occurrences et d'abondances espèces-spécifiques. Néanmoins, on peut expliquer les variations spatiales et temporelles des communautés en agrégeant les

---

<sup>5</sup> Pour décrire les relations linéaires, on peut utiliser des modèles linéaires généralisés (GLMs) suivant une distribution de Bernoulli (pour l'occurrence – ou présence) ou Poisson (pour l'abondance). Pour les relations non linéaires, on peut recourir à des modèles additifs généralisés (GAM ; [Wood, 2017](#)) ou des méthodes de machine learning (ML ; [Hastie, Tibshirani & Friedman, 2009](#), pour une vue d'ensemble des méthodes associées).

estimations des SDMs<sup>6</sup> selon des unités spatiales et temporelles arbitraires (Calabrese et al., 2014), typiquement le point de collecte et l'étendue temporelle de l'échantillonnage. Le choix de l'unité peut se faire en considération de communautés étudiées. Par exemple, si l'on désire travailler sur les communautés d'oiseaux nicheurs, on peut utiliser les données collectées durant la période de reproduction sur plusieurs années<sup>7</sup>. L'agrégation des estimations de plusieurs espèces permet de renseigner sur la composition des communautés, par exemple, la richesse spécifique, mais aussi informer sur l'équitabilité – ou *evenness* – des distributions d'abondances au sein de la communauté en incluant l'abondance relative via le calcul de l'indice de Shannon (Shannon, 1948). On peut comparer les différentes unités arbitraires selon leurs dissimilarités (Bray & Curtis, 1957) ou en comparant différentes échelles de diversité (Whittaker, 1972), à savoir la diversité locale – au sein de chaque patch ( $\alpha$ ), la diversité globale – l'ensemble des patches ( $\gamma$ ) ou leur différence par paire – *turn-over*, *i.e.* diversité ( $\beta$ ). On peut aussi considérer des dynamiques de flux entre les différents patches – *une métacommunauté* (Leibold et al., 2004).

Afin d'étudier les écosystèmes dans un cadre fonctionnel, en plus de l'abondance relative des espèces, on peut inclure des traits spécifiques comme proxy de fonctions réalisées par les espèces ou d'indicateurs de l'état des communautés. On peut étudier les dissimilarités fonctionnelles entre sites via des méthodes d'ordination (Mouchet et al., 2010; Ricotta & Bacaro, 2010) ou analyser la structure de l'espace fonctionnel (Mouillot et al., 2013). Il est aussi possible d'étudier les communautés à l'aide de traits moyens – *Community Weighted Means* (CWM ; Violle et al., 2007), afin d'étudier les variations de la diversité fonctionnelle

---

<sup>6</sup> Il est aussi possible de calculer des indices de communautés directement basés sur des données brutes collectées sur le terrain (Enemar et al., 2004; Monnet et al., 2014; Ramachandran et al., 2017; Alsila et al., 2021), mais il se peut que ceux-ci soient biaisés (Jarzyńska & Jetz, 2016). On compte aborder les biais liés aux observations dans la partie suivante *Le vivant caché par l'observation*.

<sup>7</sup> Il s'agit typiquement des données collectées par les protocoles EPOC, EPOC-ODF et STOC-EPS que l'on présente plus en détails dans la partie *Jeux de données*.

*Introduction générale*  
*caractériser ?*

*Les espèces et leurs interactions : comment les*

selon des gradients environnementaux (Gaüzère et al., 2019), par exemple, les variations des niches thermiques des communautés face aux changements climatiques (Godet, Jaffré & Devictor, 2011; Barnagaud et al., 2012; Princé & Zuckerberg, 2015).

**I.D L'observation : un filtre nécessaire à l'étude du vivant**

**I.D.i Transformation des registres et pratiques de collecte**

On constate une modification des pratiques de collecte d'information, mais aussi des perceptions humaines vis-à-vis de la faune sauvage, au cours du temps (Midgley, 1994). On peut, par exemple, noter la mention omniprésente de la chasse dans la description de l'avifaune nord-américaine réalisée par Alexander Wilson (Wilson, 1829), mais aussi, dans le cas de la Tourte voyageuse où les prélèvements dus à la chasse constituaient une source d'informations 'fiabiles' renseignant sur l'état des populations de l'espèce (Figure 10).

Petoskey, mort, par train, .....	490,000
Petoskey, vivant, par train, .....	86,400
Boyne Falls, mort, .....	47,100
Boyne Falls, vivant, .....	42,696
Petoskey, mort, par bateau, estimés, .....	110,000
Petoskey, vivant, par bateau, estimés, .....	33,640
Cheboygan, mort, par bateau, estimés, .....	108,300
Cheboygan, vivant, par bateau, estimés, .....	89,730
Autre lieux, mort et vivants, estimés, .....	<u>100,000</u>
<b>Total.....</b>	<b>1,107,866</b>

Figure 10 : Tableau dénombrant le nombre de captures de Tourte voyageuse en 1879, dans trois villes de l'État du Michigan (Petoskey, Boyne Falls, Cheboygan). A noter que l'auteur de ce tableau, E.T Martin, prenait la défense des chasseurs face aux accusations de prélèvement de Tourtes voyageuses excédant un milliard faites par le professeur Roney (Roney, 1879; Martin, 1879). Adapté de Martin 1879.

La prise en compte du bien-être animal (Tannenbaum & Bennett, 2015), ainsi que le développement de nouvelles technologies conduisent à recourir à de nouvelles pratiques de collecte (Zemanova, 2020), à savoir, le recours à l'échantillonnage d'ADN non-invasif ou non-intrusif (Lefort et al., 2022) pour renseigner sur la distribution d'espèces, principalement des poissons (Carraro et al., 2018; Minamoto, 2022) ; l'utilisation des pièges photographiques

*Introduction générale*      *L'observation : un filtre nécessaire à l'étude du vivant*  
– *camera trap* (Delisle et al., 2021) ; ou les radars (Gauthreaux & Belser, 2003; Nilsson et al., 2018) pour le suivi d'oiseaux migrateurs. De la même façon, la miniaturisation<sup>8</sup> des balises télémétriques permet d'augmenter le nombre d'espèces étudiées (Newton, 2023) tout en limitant la perte de fitness liée à la pose (Bodey et al., 2018), et le développement des unités d'enregistrement autonomes facilite le suivi acoustique de la biodiversité (Sueur et al., 2014; Gibb et al., 2019; Snyder et al., 2022).

Pour le suivi de l'avifaune, ces principaux vecteurs de transmission se concrétisent dans la réalisation d'atlas qui correspondent à des instantanés de l'ensemble des connaissances ornithologiques et présentent les états des différentes populations étudiées selon différentes échelles spatiales. On peut prendre l'exemple de l'atlas européen (Keller et al., 2020), de France métropolitaine (Issa & Muller, 2015) ou les atlas régionaux comme celui d'Auvergne-Rhône-Alpes (Clamens & LPO Auvergne, 2010). Ces atlas représentent la finalité de travaux de collecte réalisés en amont – *travaux d'atlas* – permettant de renseigner sur la distribution des zones de nidification des espèces, via des indices de nidification, ainsi que sur les effectifs reproducteurs. Au vu de l'étendue spatiale à couvrir, cette collecte d'informations fait intervenir divers acteurs, ce qui augmente considérablement le temps de collecte et d'agrégation des données. Dans le cas des atlas des espèces nicheuses de France métropolitaine, on peut noter un écart d'environ 20 ans entre chaque atlas<sup>9</sup> (Yeatman et al., 1976; Yeatman-Berthelot & Jarry, 1995; Issa & Muller, 2015).

Au fil des années, on peut constater que les atlas ont permis de couvrir des étendues spatiales plus larges (Gibbons et al., 2007; Dunn & Weston, 2008) concomitamment à une diminution du temps de collecte des travaux d'atlas (Boakes et al., 2010). Cela est principalement dû à l'essor des sciences participatives – *citizen science* – dans la collecte

---

<sup>8</sup> (Newton, 2023), note une diminution drastique du poids des transmetteurs satellites, 200 à 100g fin des années 1990 pour 5g 30 ans plus tard.

<sup>9</sup> Si l'on omet les premiers travaux de compilation de connaissance ornithologique réalisés par Mayaud en 1936 (Mayaud, Balsac & Jouard, 1936).

*Introduction générale*      *L'observation : un filtre nécessaire à l'étude du vivant*

d'informations écologiques, facilité par le déploiement et le recours à des bases de données (Porter, 2000; Hampton et al., 2013). Les sciences participatives, ou la collecte d'information scientifique par un public non professionnel (McKinley et al., 2017), permettent de renseigner sur l'état des populations (Kelling et al., 2009; Bonney et al., 2009; Matutini et al., 2021) sur de larges échelles spatiales, au travers de plateformes comme eBird ([ebird.org](http://ebird.org) ; Sullivan et al., 2009) ou iNaturalist ([inaturalist.org](http://inaturalist.org)). Malgré le double bénéfice de collecte d'information et de sensibilisation à la démarche scientifique et aux causes environnementales (Bonney et al., 2009), le recours à ces données reste difficile lorsque l'objectif de la collecte d'information n'est pas défini (Szabo, Fuller & Possingham, 2012; La Sorte et al., 2018; Bayraktarov et al., 2019). En effet, ces données sont prédisposées à être biaisées, de par un compromis à rechercher entre (i) la collecte d'une donnée de qualité, demandant un effort de collecte conséquent<sup>10</sup> et (ii) la collecte massive d'une donnée simple (Devictor, Whittaker & Beltrame, 2010)

Les principaux biais des sciences participatives non protocolées correspondent à la distribution spatiale de l'échantillonnage (Buckland & Johnston, 2017; Johnston, Matechou & Dennis, 2022), à l'hétérogénéité de l'expertise des observateurs<sup>11</sup> (Farmer et al., 2014; Schmidt et al., 2023) ou à l'effort d'échantillonnage (Callaghan et al., 2022). Tandis que les premiers biais peuvent être adressés après collecte (Johnston et al., 2019, 2020) ou en amont (Kelling et al., 2019), ces données manquent d'informations complémentaires, typiquement les distances d'observation ou les répliques temporels (King, 2014), permettant de dissocier les variations de la distribution d'abondance des espèces liées au processus écologique, de celles liées au processus d'observation (voir : *Le vivant caché par l'observation*)

---

<sup>10</sup> Un exemple de donnée coûteuse peut correspondre au protocole d'échantillonnage présenté par Morris 1955 pour l'établissement de table de survie d'insectes phytophages selon quatre zones forestières (Morris, 1955).

<sup>11</sup> On peut prendre par exemple les erreurs de type (I), ou faux-positifs, liées à la mauvaise identification des individus sur le terrain.

## Introduction générale L'observation : un filtre nécessaire à l'étude du vivant

Il est à noter que les sciences participatives représentent un nouvel aspect de la collecte d'informations (Figure 11; [Boakes et al., 2010](#)), leur utilisation dans le cas d'étude sur de larges échelles spatiales et temporelles nécessite des réflexions supplémentaires sur le 'raboutage' de données en provenance de plusieurs jeux de données différents ([Zipkin et al., 2021](#)).

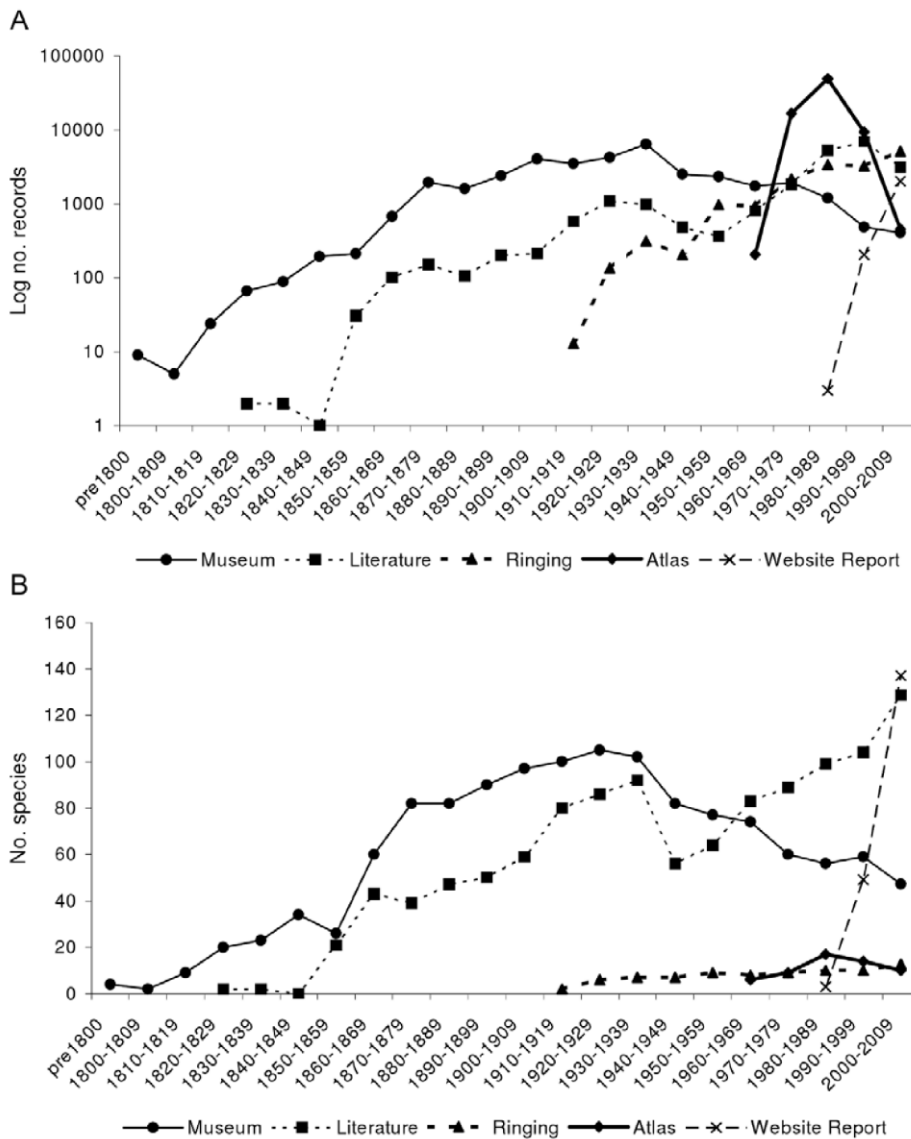


Figure 11 : Apport des différentes sources de données pour l'étude de Galliformes dans les zones biogéographiques du paléarctique et indomalais. En (A), l'évolution de l'acquisition, en échelle log, des données au fil du temps selon les différentes sources (de gauche à droite : données de collection, archives littéraires, données de capture, données d'atlas et données en provenance de base de données). En (B) L'évolution au cours du temps du nombre d'espèces renseignées selon les différentes sources. Figure tirée de Boakes et al., 2010.

I.D.ii Le suivi des effectifs d'oiseaux nicheurs

*'By going among the birds, watching them closely, comparing them carefully, and writing down, while in the field, all the characteristics of every new bird seen, - its locality, size, color, details of marking, song, food, flight, eggs, nest, and habits, - you will come easily and naturally to know the birds that are living about you. The first law of field work is **exact observation**, but not only are you more likely to observe accurately if what you see is put in black and white, but you will find it much easier to identify the birds from your notes than from your memory.'*, Bailey, 1898 – Birds through an opera glass

Les oiseaux sont un modèle d'étude pertinent pour la production de bioindicateurs, de par : (i) leur observation facile (Fiedler, 2009) ; (ii) la connaissance approfondie de leur écologie et de leur cycle de vie, en partie initiée par les observations de Gilbert White à la fin du 18e siècle (White, 1789) lorsque l'activité naturaliste était encore considérée comme un passe-temps inhabituel (Moss, 2004); et (iii) leur présence sur divers niveaux trophiques qui permet de mesurer les effets des changements globaux comme par exemple l'accumulation de certains pesticides – le dichlorodiphényltrichloréthane (DDT) chez les oiseaux de proie (Bowerman et al., 1995; Hellou, Lebeuf & Rudi, 2013).

C'est dans l'optique d'étudier les variations temporelles des populations d'oiseaux, ainsi que l'effet des activités anthropiques (Butcher et al., 1987; Sauer et al., 2017), que se sont développés les suivis d'effectifs d'oiseaux nicheurs – *Breeding Bird Surveys* (BBS ; Butcher et al., 1987; Sauer et al., 2017). On note une volonté de structurer la collecte d'informations dès 1932 avec la création du British Trust for Ornithology (BTO ; Moss, 2004), suivie du lancement de plusieurs programmes de suivis, notamment les programmes américain<sup>12</sup> en 1965 (Butcher et al., 1987), puis français en 1989 avec le Suivi Temporel des

---

<sup>12</sup> On peut aussi noter l'initiative du Christmas Bird Count lancée en décembre 1900 par la revue Bird Lore (Chapman, 1901).

*Introduction générale*      *L'observation : un filtre nécessaire à l'étude du vivant*  
Oiseaux Communs (STOC; [Julliard & Jiguet, 2002](#); [Fontaine et al., 2020](#)) et britannique<sup>13</sup> en  
1994 ([Gregory et al., 1996](#)).

La collaboration de multiples observateurs a rendu nécessaire la standardisation de la collecte de données ornithologiques, par exemple, vis-à-vis de la classification des habitats ([Crick, 1992](#)), en amont des données collectées par images satellites ([Luque et al., 2018](#)). L'objectif est de prendre en compte les divers facteurs affectant les variations du nombre d'individus détectés (Figure 12; [Emlen, 1971](#)) et d'assurer une estimation relativement robuste en appliquant un coefficient de correction entre le nombre d'individus détectés et le nombre d'individus estimés de la population ([Emlen, 1971](#)). Pour cela, différentes méthodes d'estimation peuvent être utilisées, à savoir le décompte de nids ([Emlen, 1971](#)), la cartographie de territoires – *territory mapping* ([Verner, 1985](#)), la capture-marquage-recapture ([Nichols et al., 1981](#)) le décompte d'individus par linéaire (Burnham 1980) ou la réalisation de points d'écoute, avec plusieurs variantes, comme les IPA (Indices Ponctuels d'Abondance) et les EFP (Echantillonnage Fréquentiel Progressif ; [Blondel, Ferry & Frochot, 1981](#)). En effet, les données non protocolées, ou 'opportunistes', sont plus difficilement exploitables, malgré leur abondance, en raison de l'impossibilité de quantifier l'effort d'observation et les données d'absence ([Bonney et al., 2009](#); [Johnston et al., 2019](#)). La mise en place progressive de règles régissant la collecte de données, concernant par exemple la restriction de la période de collecte, l'échantillonnage aléatoire, la quantification de la zone couverte ou la limitation du temps de collecte a eu pour objectif d'homogénéiser les données collectées sur le terrain, afin d'accroître leur comparabilité et leur qualité ([Kelling et al., 2009](#)) et permettre l'utilisation d'outils statistiques de plus en plus exigeants ([Jeliazkov et al., 2022](#)).

---

<sup>13</sup> Avant cela, les espèces britanniques étaient suivies par le Common Bird Census lancé en 1959. Le BBS britannique a vu le jour afin de surmonter des problèmes liés au manque de représentativité du territoire provoqué par une collecte d'information focalisée vers le sud-ouest du territoire et le comportement de collecte des observateurs, une surreprésentation des territoires agricoles ([Gregory et al., 1996](#)).

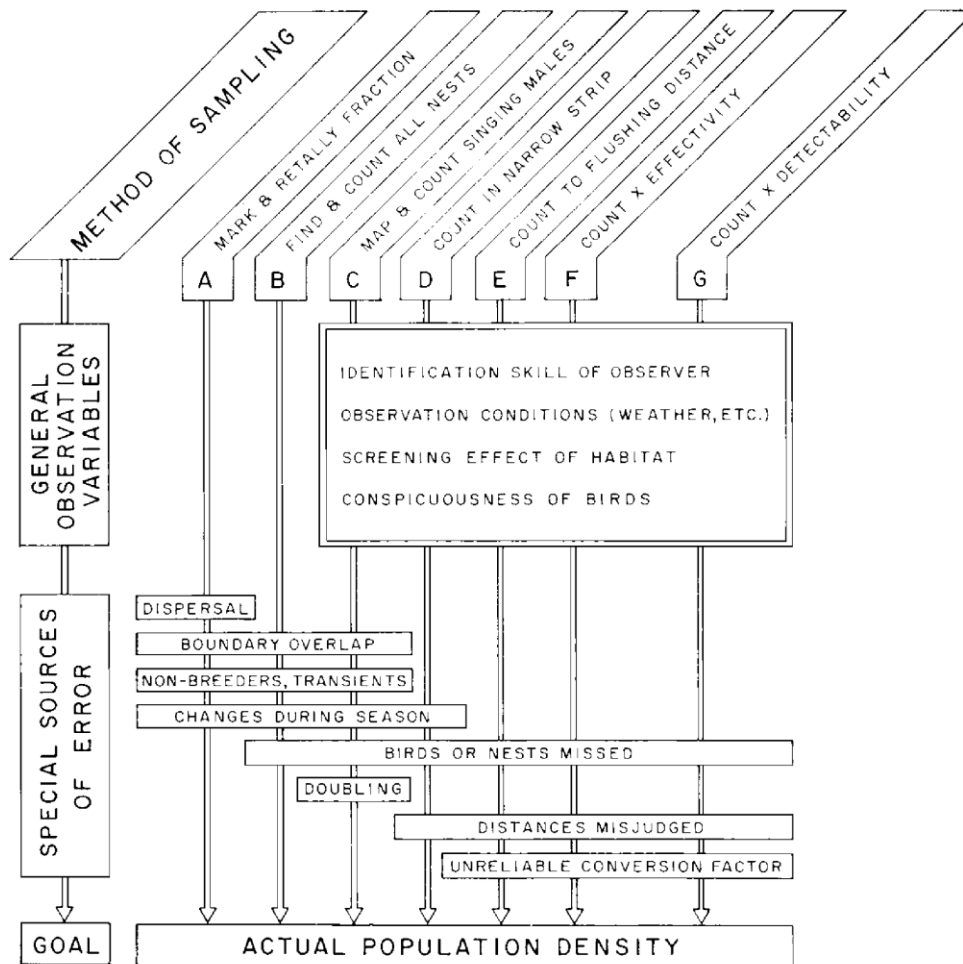


Figure 12 : Illustration des facteurs limitant l'efficacité et la précision des méthodes de décompte d'oiseaux terrestres ne se rassemblant pas en groupes. En haut décrites par des lettres majuscules, les différentes pratiques de collecte d'informations. (A) baguage d'une partie de la population; (B) décompte de nids; (C) cartographie des territoires de mâles chanteurs; (D) décompte selon des surfaces (linéaires ou points); (E) décompte en renseignant les distances d'observation; (F) décompte en intégrant la performance des observateurs; (G) décompte en intégrant la détectabilité de l'espèce cible. Figure tirée d'Emlen 1971.

La standardisation de la collecte de données repose sur l'hypothèse forte que le nombre d'individus détectés représente une part constante du nombre total d'individus (Thompson, 2002). Cela n'est pas forcément vrai (Burnham, 1981) au vu des disparités liées aux habitats, aux espèces, aux observateurs ou au temps de réalisation de la liste (Figure 12; Pacifici, Simons & Pollock, 2008; Guillerá-Arroita, 2017). Il est crucial de tenir compte de ces variations pour obtenir des estimations d'abondance ou d'occupation fiables (Nichols, Thomas & Conn, 2009).

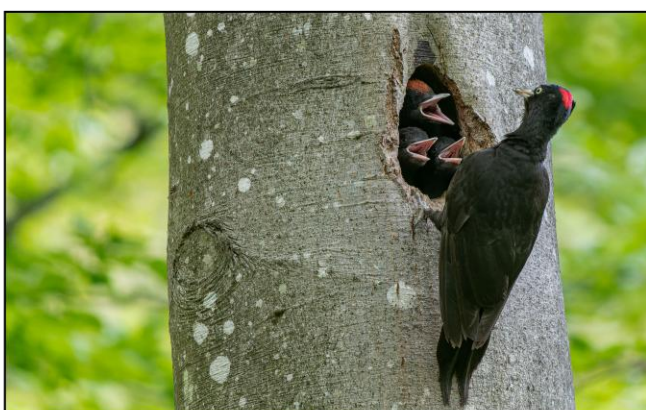


Illustration : Signes et révélations. Crédit photo Pic noir (*Dryocopus martius*) : Christian de Muizon et capture d'écran de 'The Witness', Thekla Inc., 2016

I.D.iii Le vivant caché par l'observation

*'But as with the appearance of the sun the visible world makes its appearance, so at one stroke does the understanding through its one simple function convert the dull meaningless sensation into perception. What the eye, the ear, or the hand experiences is not perception; it is mere data.'*, Arthur Schopenhauer, Section 4, p12., *The World As Will and Representation*. Vol. 1. 1966.

Sans devoir remettre en question l'ensemble de nos perceptions et concepts nous permettant d'interpréter notre monde, on est en droit de douter de l'observation et se poser la question suivante : face au processus étudié, à quoi correspondent les données issues d'observation ? Se rapportent-elles parfaitement au processus, ou uniquement en partie ? Et finalement, si la correspondance est partielle, comment est régie cette scission ?

Dans nos travaux, le processus étudié correspond à la distribution spatiale de l'abondance des individus décrite par un ensemble de relations linéaires aux conditions environnementales<sup>14</sup>. Cette quantité d'individus peut se baser sur (i) des estimations absolues<sup>15</sup>, *i.e.* la quantité d'individus par surface ; ou (ii) des estimations relatives (Southwood & Henderson, 2000). Le recours à l'une ou l'autre de ces estimations est en partie dû à l'échelle spatiale de l'étude et au coût (humain et financier) de la collecte de l'information (Morris, 1955, 1960; Thompson, White & Gowan, 1998). Mais ce choix est aussi lié à la question de recherche: on favorise les estimations absolues pour déterminer l'état de conservation des populations (Critères A,C et D – IUCN, 2012) et leurs processus démographiques (Thompson, White & Gowan, 1998; Pollock, Nichols & Karanth, 2012). Les

---

<sup>14</sup> Voir les parties 'Hierarchical Distance Sampling : L'estimation d'abondance à partir de données de distance' et 'Environnement et modèles d'études' pour plus de détails sur ce point.

<sup>15</sup> Les estimations absolues peuvent aussi être obtenues par des proxys, par exemple on peut dénombrer la quantité d'individus selon une ressource utilisée par l'espèce étudiée qui est plus facilement dénombrable. Morris 1955 prend comme exemple le cas d'insectes phyllophages dénombrés à partir du nombre de bourgeons.

*Introduction générale* L'observation : un filtre nécessaire à l'étude du vivant  
indices relatifs sont favorisés pour l'étude des distributions d'espèces à large échelle, l'évaluation des impacts environnementaux ou renseignés sur l'activité des espèces (Southwood & Henderson, 2000).

Les objectifs de la biogéographie de la conservation, notamment l'étude des processus à large échelle spatiale et temporelle (Whittaker et al., 2005) contribuent à favoriser le recours aux indices relatifs. Cependant, ces indices reposent sur l'hypothèse que la variation de l'indice est uniquement liée à des variations de la variable d'intérêt (Figure 13; Thompson, White & Gowan, 1998).

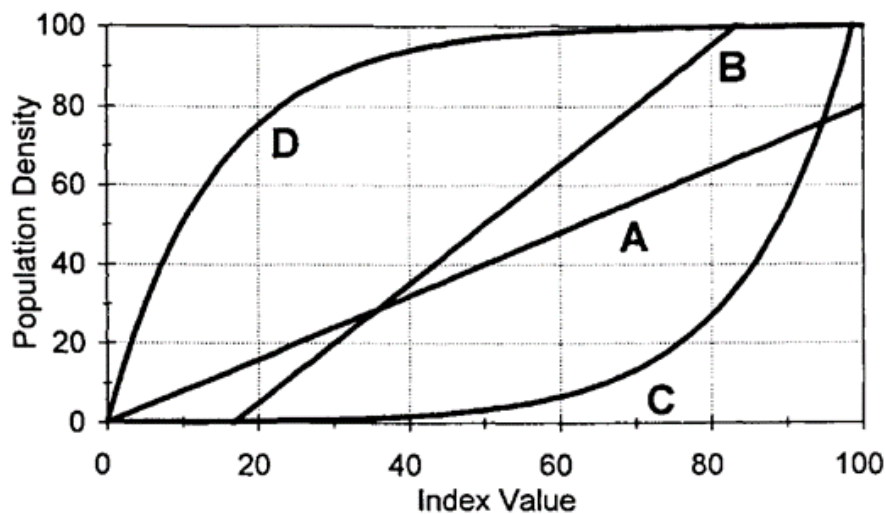


Figure 13 : Représentation de différents types de relations entre un indice relatif (en abscisse) et un paramètre d'intérêt, ici la densité d'individus (en ordonnée). La relation entre les deux indices peut être soit linéaire sans biais, *i.e.* intercept à 0 (A) ; linéaire avec un biais constant (B) ; ou des relations non linéaires (C) et (D), où les variations de densité 'réelle' n'induisent pas de changements proportionnels de l'indice relatif. Pour ces deux derniers cas, (Thompson, White & Gowan, 1998) prennent l'exemple de données issues de cartographie de territoire (Verner, 1985). On s'attend à ce que les individus délimitent mieux leur territoire, *e.g.* augmentation de la fréquence des chants, à de fortes densités facilitant leur détection (C), ou inversement, pour une espèce avec peu d'individus territoriaux (D). Thompson et al., 1998.

Dans le cadre des décomptes d'oiseaux, les estimations basées sur des points d'écoute, où l'on compte le nombre d'individus détectés à la vue et/ou l'ouïe, reviennent à

*Introduction générale*      *L'observation : un filtre nécessaire à l'étude du vivant*

considérer que le nombre d'individus détectés correspond à une proportion constante du nombre total d'individus (Burnham, 1981; Thompson, 2002). Cependant, la relation entre l'indice relatif et le paramètre de population d'intérêt peut avoir plusieurs formes (Figure 13), elle peut soit (i) répondre convenablement à l'hypothèse (Figure 13A) ;(ii) répondre convenablement à l'hypothèse, mais présenter un biais<sup>16</sup> (Figure 13B) ; ou (iii) la relation n'est pas constante (Figure 13C et Figure 13D).

Or, la variation du nombre d'individus observés peut être causée par deux sources de variations : (i) les variations liées à l'habitat, *i.e.* on s'attend à ce qu'une espèce soit présente en plus grande abondance dans des habitats correspondant à sa niche écologique – *processus écologique* ; (ii) les variations liées à l'observation, *i.e.* il est possible que l'ensemble des individus ne soit pas présent et/ou détecté lors de la réalisation d'un point d'écoute – *processus d'observation* (Royle & Dorazio, 2008; Kéry & Royle, 2015). Les variations liées au processus écologique sont prises en compte par l'échantillonnage des habitats (Figure 14) et l'hypothèse de la relation espèce-habitat (par la formulation du modèle). Par contre, les variations liées aux processus d'observations peuvent être multiples (Kellner & Swihart, 2014), et incluent, par exemple, le jour de réalisation du point d'écoute (Schmidt, McIntyre & MacCluskie, 2013), l'expertise des observateurs (Kendall, Peterjohn & Sauer, 1996; Farmer, Leonard & Horn, 2012; Schmidt et al., 2023) ou encore les traits morphométriques et comportementaux des espèces (Johnston et al., 2014; Sóllymos et al., 2018; Morelli et al., 2022).

---

<sup>16</sup> A noter, que les relations A et B représentent les erreurs de type (II, faux-négatifs) et (I, faux-positifs). Pour la relation A, on observe que celle-ci n'est pas de forme  $y = x$ , supposant que le nombre d'individus dénombrés par l'indice relatif est inférieur à la densité attendue. Pour la relation B, on voit que l'intercept n'est pas à 0 suggérant que l'indice relatif décompte des individus absents.

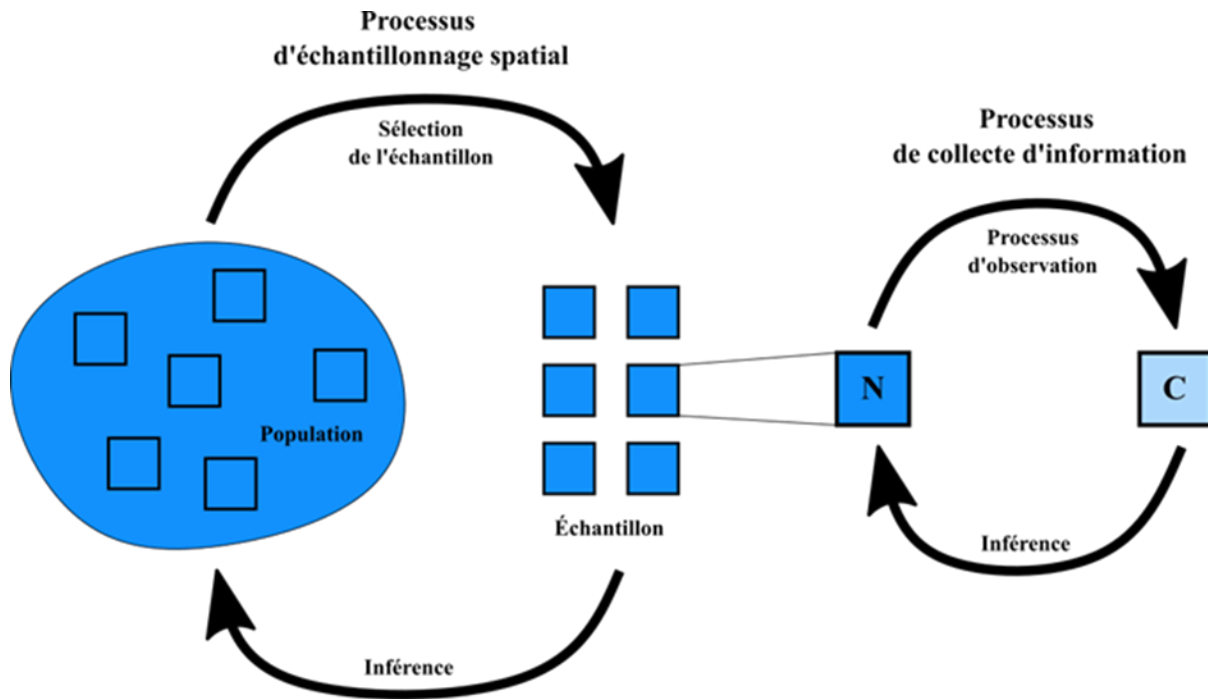


Figure 14 : Représentation des processus stochastiques sous-jacents à l'estimation d'abondance. L'échantillonnage spatial permet d'estimer la quantité d'individus (N) qui conditionne le nombre d'individus détectés (C) lors de la réalisation de la liste. Adapté de la figure 1.4 - Kéry et al., 2015.

Ces deux processus peuvent être représentés par la succession de plusieurs modèles de probabilité ordonnée par des paramètres conditionnellement dépendants les uns des autres, des *modèles hiérarchiques* (Royle & Dorazio, 2008; Kéry & Royle, 2015). Sur la Figure 14, on peut constater que la quantité d'individus détectée (C) dépend du nombre d'individus disponibles (N), lui-même dépendant de paramètres de population décrivant le nombre d'individus et leur répartition spatiale. Les modèles hiérarchiques permettent de décrire un système complexe (Figure 14) en un ensemble de sous-systèmes plus simples permettant d'inférer sur les paramètres de la population à partir des données de comptages tout en partitionnant les incertitudes provenant des deux processus (Hobbs & Hooten, 2015).

Nichols, Thomas & Conn, 2009 proposent un cadre conceptuel qui décrit ces deux processus au travers de quatre probabilités :  $P_s$ , la probabilité que l'échantillonnage recoupe les habitats favorables à l'espèce ;  $P_p$ , la probabilité que les individus soient présents lors des

*Introduction générale*      *L'observation : un filtre nécessaire à l'étude du vivant*

visites des observateurs ;  $P_a$ , la probabilité que les individus se rendent disponibles à l'observation, *i.e.* émettent un chant, ou s'exposent sur des branches ; et  $P_d$ , la probabilité que les individus disponibles soient détectés par les observateurs. A noter que dans cette description des processus,  $P_s$  s'apparente au processus écologique et  $P_p$  décomposé en  $P_a$  et  $P_d$  au processus d'observation. L'estimation des différentes composantes du processus d'observation requiert différentes approches de collecte de données (Pollock, 1982; Southwood & Henderson, 2000; King, 2014), notamment les répliques temporels (Chandler, Royle & King, 2011), l'acquisition de données de distance d'observation (Royle, Dawson & Bates, 2004) ou le recours à plusieurs observateurs (Nichols et al., 2000).

Si l'on retranscrit la représentation dépeinte par Nichols en termes de nombre d'individus selon un plan d'échantillonnage composé de trois visites, on obtient la figure suivante (Figure 15) où la quantité totale d'individus ( $M$ ) peut être en partie disponible lors d'un ou de multiples répliques temporels ( $N_1, N_2, N_3$ ) et qu'au sein de ces différents répliques, certains individus peuvent être manqués par les observateurs. Nous comptons décrire le processus plus en détail au sein de la partie *Hierarchical Distance Sampling : L'estimation d'abondance à partir de données de distance*.

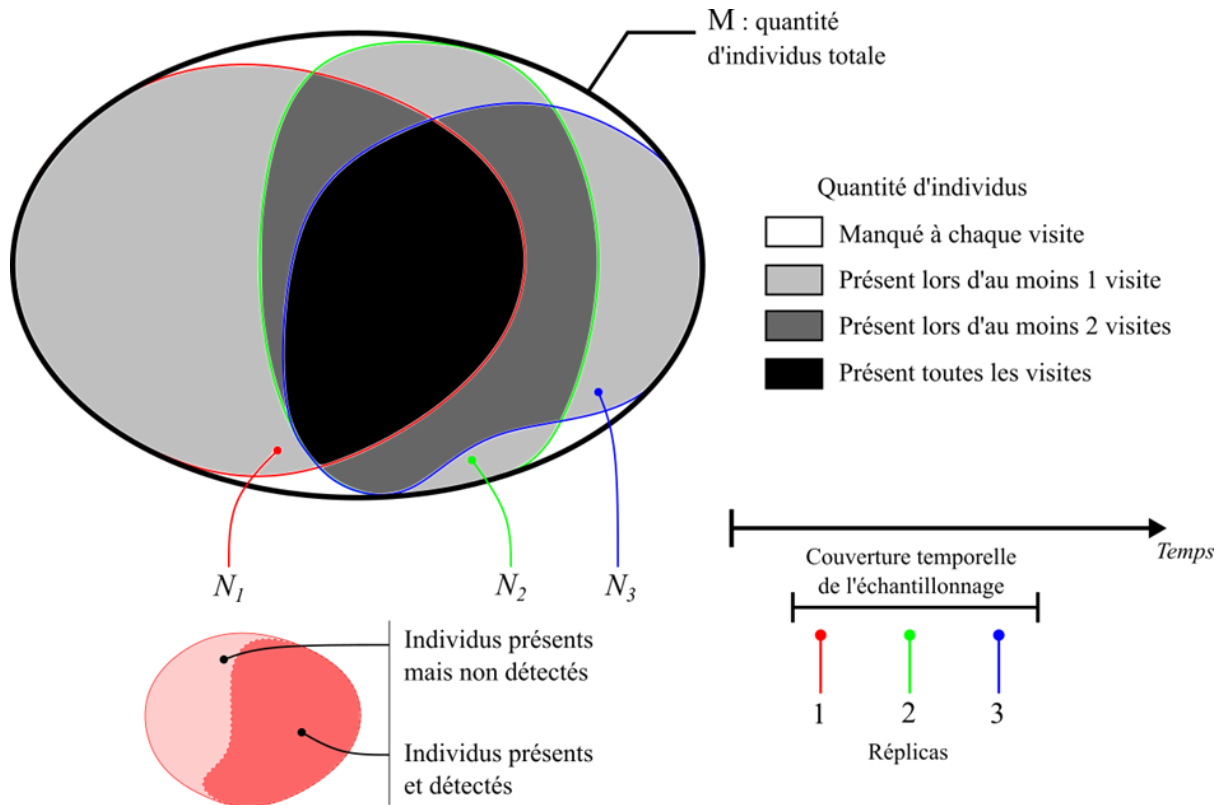


Figure 15 : Visualisation des différentes probabilités du processus d'observation selon la quantité d'individus détectée.  $M$  représente la quantité d'individus totale sur l'ensemble de sites de collecte.  $N_1$ ,  $N_2$  et  $N_3$  correspondent aux quantités d'individus qui se rendent disponibles à l'observation durant les trois réplicas. Au sein même d'un réplica, on s'attend à ce qu'une partie des individus ne soit pas détectée. Durant les différentes visites, certains individus peuvent être présents et disponibles sur les sites plus d'une fois (dégradé de gris) ou totalement manqués (blanc).

## I.E Intérêt de conservation : la place du commun

‘The global biodiversity crisis is not simply one of species loss.’, Gaston et al., 2018

La pratique de la conservation s’oppose à des difficultés de financement de ses actions. A la fin du 20<sup>e</sup> siècle, on estimait l’ensemble des coûts des mesures de conservation à six milliards USD contre trente-trois trilliards<sup>17</sup> USD de gains potentiels liés aux effets directs et indirects des services écosystémiques (Mace, Possingham & Leader-Williams, 2008). Bien que les montants alloués aux mesures de conservation aient augmenté ces 20 dernières années, entre 124 et 143 milliards USD par an, on reste encore loin des estimations de 722 à 967 milliards USD par an afin d’enrayer le déclin de la biodiversité mondiale d’ici 2030 (Deutz et al., 2020).

Ce manque de financement favorise une priorisation des mesures de conservation et contraint à choisir entre des mesures de conservation basées sur les écosystèmes ou basées sur les espèces (Mace, Possingham & Leader-Williams, 2008). Tandis que la première approche permet d’intégrer les interactions des écosystèmes dans l’établissement des mesures de conservation (Rosenberg & McLeod, 2005), la seconde permet de plus facilement suivre l’efficacité des mesures et/ou renseigner sur l’état des milieux auprès d’un plus large public (Mace, Possingham & Leader-Williams, 2008).

En effet, on peut noter l’apparition de concepts comme celui des espèces clés – *keystone species*, et des espèces parapluies – *umbrella species*, qui décrivent, pour l’une, des espèces ayant un rôle essentiel aux fonctions des écosystèmes, et pour l’autre, des espèces dont la conservation permettrait d’assurer la pérennité d’autres espèces (Simberloff, 1998; Mace, Possingham & Leader-Williams, 2008). D’autres notions promeuvent des espèces à des fins plus stratégiques, comme les espèces étendards – *flagship species*, pour sensibiliser

---

<sup>17</sup> 20 ans plus tard, ces estimations s’élèvent à 44 trilliards USD (Herweijer et al., 2020)

l'attention du public, ou les espèces indicatrices – *indicator species*, synthétisant la réponse aux changements globaux d'un ensemble d'espèces (Mace, Possingham & Leader-Williams, 2008). Les ressources et l'attention peuvent être concentrées sur un sous-ensemble d'espèces pour des raisons pragmatiques et valables (Bottrill et al., 2008, 2009), à savoir mettre en œuvre des actions de conservation envers les espèces en état de conservation critique et propices aux vortex d'extinction (Gilpin & Soulé, 1986), mais le recours à cette priorisation nécessite un questionnement éthique (Wilson & Law, 2016).

Dans le cas de la surveillance des effectifs d'oiseaux nicheurs de France métropolitaine, ce 'délaissement' se traduit par la production d'estimations d'effectifs plus fiables pour les espèces en voie d'extinction et/ou exploitées. En effet, ces espèces sont sujettes à des suivis exhaustifs (Quaintenne et al., 2020), des Plans Nationaux d'Actions (PNA ; Challear & Lavarde, 2014) ou des suivis par des programmes protocolés, notamment les programmes LIMAT (Limicoles et Anatidés; Aubry et al., 2023) ACT (Alaudidés, Colombidés et Turdidés) et 'Flash' qui évaluent les tendances démographiques et effectifs des populations nicheuses et hivernantes d'espèces chassables (Réseau "Oiseaux de passage," 2017).

Lors du rapportage de la Directive Oiseaux européen (Directive 2009/147/CE du Parlement européen et du Conseil du 30 novembre 2009 concernant la conservation des oiseaux sauvages, 2009), qui demande à l'ensemble des états membres de renseigner sur l'état des populations d'oiseaux nichant sur leur territoire, sur les 292 espèces d'oiseaux nicheuses en France métropolitaine, 65 %, principalement les espèces communes, ont des estimations d'effectifs qualifiés comme 'moyennes' (Comolet-Tirman et al., 2022) reflétant un état des connaissances plus ou moins bon associé à des données quantitatives incomplètes ou périmées (Comolet-Tirman et al., 2015). Ce constat est le témoin d'une pratique de la

conservation principalement focalisée sur le sauvetage d'espèces en voie d'extinction et non sur les phénomènes d'épuisement global des populations (Gaston & Fuller, 2008), et ce malgré l'importance des espèces communes dans la stabilité des écosystèmes et les services écosystémiques qu'elles peuvent apporter (Figure 16; Gaston et al., 2018).

C'est dans ce contexte que la Ligue pour la Protection des Oiseaux (LPO), à la suite et en collaboration avec le MNHN, en tant que membres du groupe de coordination nationale 'Directive oiseaux', a pris en charge la mise à disposition de jeux de données assurant le suivi d'abondance et de distribution des oiseaux nicheurs. Ces jeux de données proviennent de plusieurs protocoles de suivis, tels que le STOC, l'Estimation des Populations d'Oiseaux Communs dans le cadre de l'atlas des Oiseaux de France (EPOC-ODF) et EPOC<sup>18</sup>, mais aussi les données opportunistes collectées via le portail d'observation Faune-France ([faune-france.org](http://faune-france.org)).

---

<sup>18</sup> L'ensemble de ces protocoles sont présentés plus en détails dans les sections qui suivent (cf. *Jeux de données*)

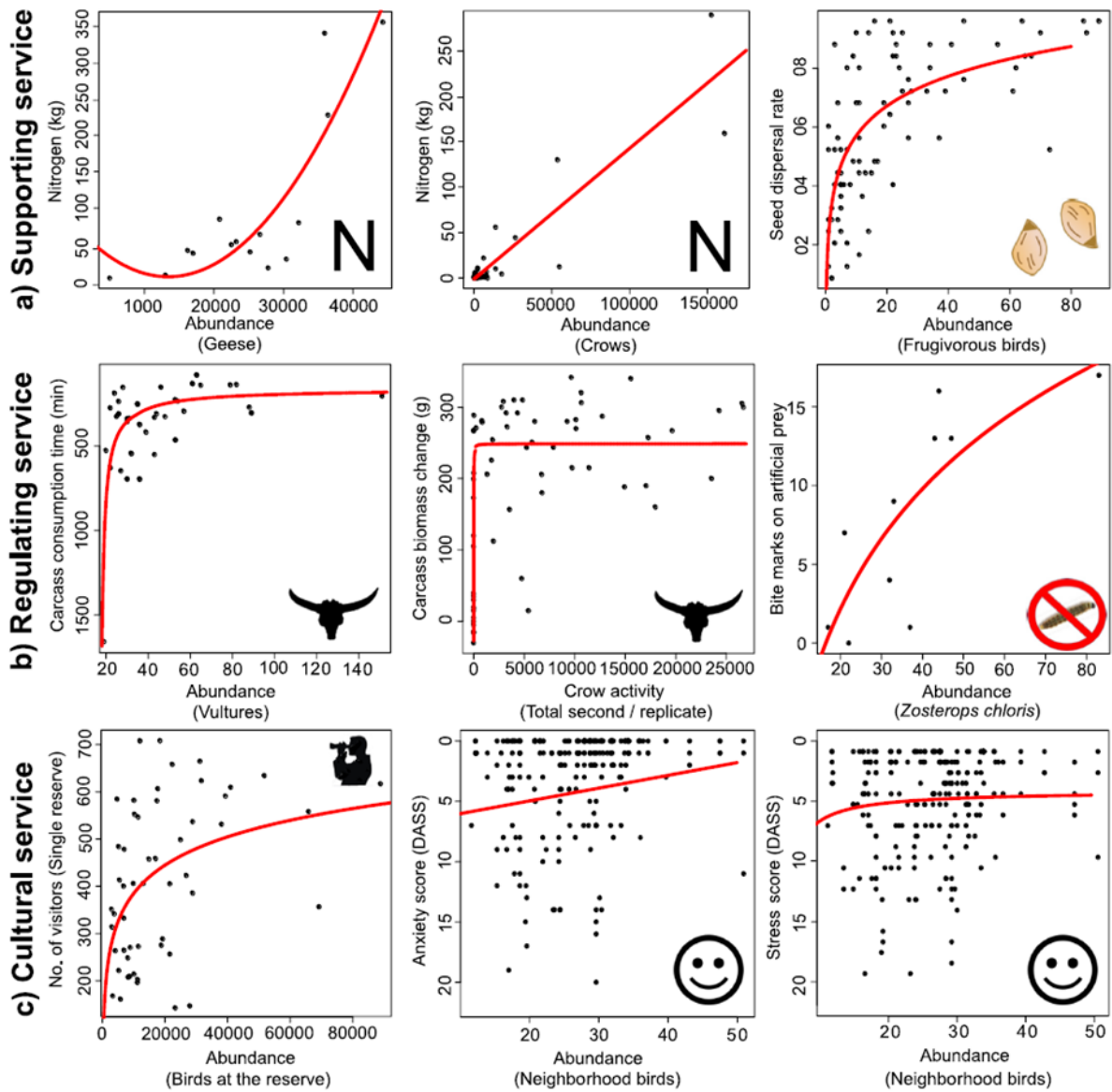


Figure 16 : Ensemble d'études montrant les relations entre l'abondance des oiseaux et les services écosystémiques. Chaque ligne représente un type de service écosystémique spécifique (voir figure A du Millenium Ecosystem Assessment – MEA 2005). De haut en bas et de gauche à droite, (a) services d'approvisionnement (transport de nutriments par les anatidés et les corvidés ; dispersion des graines) ; (b) services de régulation (équarrissage naturel par les vautours et corvidés ; gestion des ravageurs de cultures) ; et (c) services culturels (tourisme ; diminution des niveaux d'anxiété et de stress). Figure tirée de Gaston et al., 2018

## I.F Objectifs et questions de recherches

De par son encadrement, cette thèse CIFRE (Convention Industrielle de Formation par la REcherche) a un double objectif : (i) l'utilisation et le développement de modèles permettant d'estimer plus finement les tailles de populations et l'évaluation des communautés d'oiseaux, et (ii) la production de connaissances permettant d'évaluer au mieux les populations d'oiseaux communs de France métropolitaine afin de répondre aux demandes de la Commission Européenne. Cette thèse s'organise en trois chapitres autour de ces deux axes. Je me suis focalisé sur l'étude des variations spatiales d'abondances des populations d'oiseaux en France métropolitaine en période de reproduction en prenant en compte les variations dues à la collecte de données – *le processus d'observation*.

**Pour le premier chapitre**, j'ai travaillé sur le développement d'une démarche permettant d'estimer les effectifs nicheurs d'un ensemble d'espèces communes à partir des données du protocole structuré EPOC-ODF (Estimation des tailles de Populations d'Oiseaux Communs – Oiseaux De France). Ce protocole a été conçu pour collecter des données ornithologiques de grande qualité, au travers de répliques temporels. L'ensemble de ces données a permis le calibrage de modèles hiérarchiques de Distance Sampling (Chandler, Royle & King, 2011) permettant de dissocier les variations d'abondances dues à l'habitat, *processus écologique*, de celles provoquées par la collecte, *processus d'observation*. J'ai comparé ces nouvelles estimations à celles du précédent atlas 'Atlas des oiseaux de France métropolitaine – 2009/2012' (Issa & Muller, 2015) obtenues dans un contexte de collecte partielle sans recours à des modèles statistiques. Comme vu précédemment (cf : *Le vivant caché par l'observation*), l'omission du processus d'observation tend à biaiser les estimations absolues (Numminen et al., 2023), l'objectif principal de ce chapitre tente de répondre à la question suivante : *Est-ce que les populations d'oiseaux nicheurs étaient correctement estimées lors du précédent atlas ?*

Les données de sciences participatives sont confrontées à un compromis entre (i) la collecte d'informations peu, voire non structurées, en grande quantité et sur de larges étendues spatiales et (ii) la collecte d'informations structurées, moins abondantes (Devictor, Whittaker & Beltrame, 2010). Dans le premier chapitre, nous avons utilisé les données structurées d'EPOC-ODF. La LPO avait lancé un autre protocole de science participative, l'EPOC. Contrairement à l'EPOC-ODF, le protocole EPOC ne reposait ni sur une présélection des sites de collecte ni sur la réalisation de répliques temporels. Ainsi, pour une période identique (2021-2023), le protocole EPOC a permis de collecter trois fois plus de données que l'EPOC-ODF.

Afin de chercher à valoriser ces données disponibles, **pour le second chapitre**, j'ai travaillé au développement d'un modèle intégré (Fletcher et al., 2019; Zipkin et al., 2021), d'estimation des tailles de populations d'oiseaux communs nicheurs en utilisant conjointement les données collectées par les deux protocoles. Je me suis basé sur le modèle de distance sampling intégré – Integrated Distance Sampling (IDS ; Kéry et al., 2024), en me focalisant sur l'estimation de la disponibilité des individus lors de l'observation (voir : *Le vivant caché par l'observation*). Cette information est accessible au travers des répliques temporels des EPOC-ODF, qui n'existent pas pour les données EPOC. Au travers de cette intégration de données, je tente de répondre à la question suivante: *Les modèles de données intégrés permettent-ils d'améliorer les estimations d'effectifs nicheurs ? Représentent-ils des moyens concrets d'amélioration des estimations d'effectifs nicheurs ?*

Dans le cadre **du troisième chapitre** de cette thèse, je m'intéresse à la caractérisation des communautés d'oiseaux nicheurs au travers d'indices taxonomiques, phylogénétiques et fonctionnels (Maurer & McGill, 2010). Les communautés sont composées de multiples espèces avec des spécificités biologiques et comportementales qui influent de façon

## *Introduction générale*

## *Objectifs et questions de recherches*

hétérogène sur leur détectabilité (Iknayan et al., 2014). Je me suis intéressé ici aux données collectées par le protocole de sciences participatives structuré STOC-EPS (Suivi Temporel des Oiseaux Communs par Échantillonnage Ponctuel Simple) afin de comparer des indices estimés avec une prise en compte ou non du processus d'observation, en tentant de répondre à la question suivante: *L'omission du processus d'observation dans l'étude des communautés, via l'estimation de traits moyen, produit-il des estimations biaisées ?*



## II Jeux de données et Méthodes

---

### II.A Jeux de données

#### II.A.i STOC-EPS (Suivi Temporel des Oiseaux Communs par Échantillonnage Ponctuel Simple)

Le protocole STOC-EPS (STOC ou FBBS : French Breeding Bird Survey) est un dispositif de suivi de l'avifaune française, reposant sur des données collectées par des observateurs volontaires expérimentés, lancé en 1989, sous la responsabilité de Christian Vansteewegen du Centre de Recherche sur la Biologie des Populations d'Oiseaux (CRBPO) au Muséum National d'Histoire Naturelle (MNHN ; [Julliard & Jiguet, 2002](#)).

Dès son origine, le protocole STOC-EPS fut conçu comme étant un suivi complémentaire à d'autres dispositifs, notamment le protocole STOC-capture, qui par la capture et le marquage d'individus vise à quantifier les variations des paramètres démographiques, mais nécessite une capture au filet, des compétences et des autorisations de capture d'oiseaux. Par sa simplicité relative de mise en place, le STOC-EPS a un potentiel de collecte de données d'observation standardisées sur l'avifaune française ([Julliard & Jiguet, 2002](#)). Le dispositif fut revisité en 2001, avec pour volonté d'harmoniser la collecte d'information à celle d'autres dispositifs du réseau européen PanEuropeanCommon Bird Monitoring Scheme (PECBMS). Ces modifications avaient pour objectif d'assurer une distribution homogène des habitats en attribuant aléatoirement une maille de 2x2km à chaque observateur, selon sa localité (à moins de 10 km; [Jiguet et al., 2010](#)). Hormis de légères modifications en 2015 permettant de standardiser la collecte (notamment au sujet de la géolocalisation des sites de collecte), la forme actuelle du protocole correspond à celle de 2001.

Il est demandé aux observateurs de prospecter 10 sites – points STOC – au sein de ces mailles. Lors de la sélection des sites, une attention particulière doit être portée sur la représentativité des habitats au sein de la maille 2x2km, ainsi que sur un espacement des points STOC d’au moins 300m les uns des autres (Figure 17; Jiguet et al., 2012).

Le protocole ciblant principalement la période de reproduction, il est demandé de prospecter les points STOC deux fois durant cette période (1er passage : du 01/04 au 08/05; 2eme passage : du 09/05 au 15/06). Depuis trois ans, les observateurs peuvent aussi réaliser un passage “précoce” du 01/03 au 31/03.

Le protocole se base sur des points d'écoute de 5 minutes, pendant lesquels les observateurs doivent identifier et compter le nombre d'individus contactés, par l'ouïe ou la vue, autour d'eux (Figure 17), en estimant la distance d'observation (Observateurs - Individus détectés) selon 4 classes de distance (0-25m ; 25m-100m; 100m-200m et >200m) et en notant si les individus contactés étaient en vol ou stationnaires, la catégorie 'en vol' étant considérée comme une classe de distance.

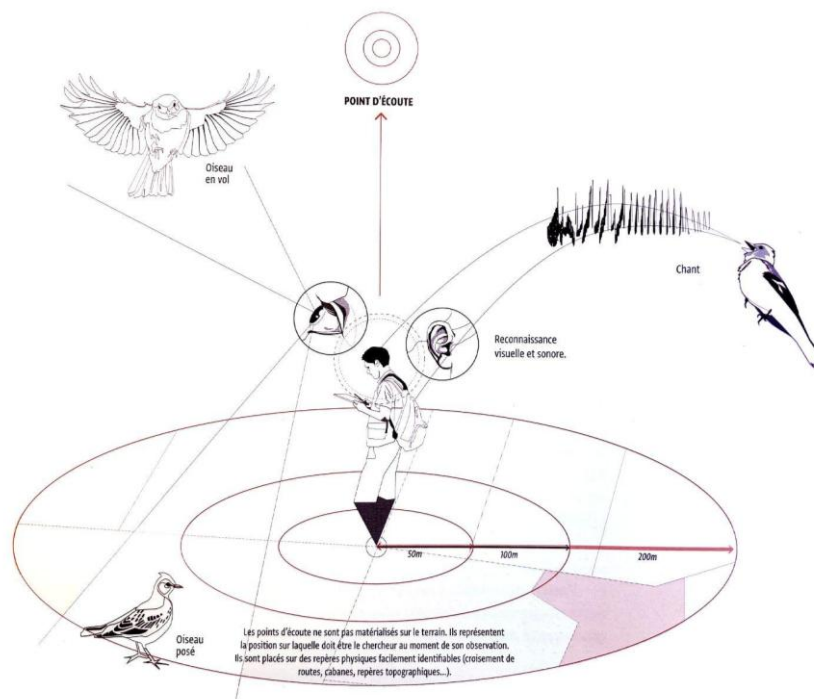


Figure 17 : Représentation schématique de la réalisation d'un point d'écoute STOC. La détection peut être auditive comme visuelle. Durant le point d'écoute, les observateurs sont amenés à renseigner la distance d'observation estimée dans quatre classes de distances (0-25m; 25-100m; 100-200m; +200m). Crédit : Dessin - Axelle Grégoire - Extrait de l'article "Étudier le vivant à l'échelle du paysage"- Revue Billebaude n°13 "Affronter la sixième extinction" -Automne-hiver 2018 - Fondation François Sommer (Paris) et Editions Glénat (Grenoble).

## II.A.ii EPOC (Estimation des Populations d'Oiseaux Communs)

Le programme EPOC est un programme de science participative lancé en 2017 par la LPO ayant pour objectif d'être opérationnelle pour estimer les tailles des populations d'oiseaux communs en France. Comme le STOC-EPS, le protocole EPOC se base sur la participation volontaire d'observateurs afin de collecter des données de comptage lors de la période de reproduction. Cependant, à l'inverse de ce premier, le protocole EPOC ne requiert pas la sélection d'un site en particulier.

Le protocole EPOC se base sur des points d'écoute de 5 minutes, où les observateurs notent la distance d'observation entre eux et les individus contactés. Afin d'aider à la collecte d'informations, la LPO utilise une application mobile ([NaturaList](#)) correspondant à la version mobile de la plateforme [Faune-France](#). L'objectif est de faciliter la collecte d'informations, en permettant la visualisation des habitats environnants via un fond de cartes, ainsi que la qualité de la donnée collectée (distance calculée à posteriori et non plus basée sur les classes de distances du STOC-EPS).

Le protocole EPOC est un protocole dit "semi-structuré", l'idée étant de collecter une donnée standardisée sous la forme de liste complète de 5 minutes, tout en laissant une liberté dans la sélection des sites aux observateurs (Figure 18). Ce protocole a pour intention de tirer profit de l'engouement pour la collecte de données naturalistes par les bases de données en ligne ([ebirg.org](#) ; [inaturalist.org](#) ; [faune-france.org](#)).

### II.A.iii EPOC-ODF : Un protocole structuré pour l'estimation des populations d'oiseaux communs

Le protocole EPOC-ODF est un dispositif lancé en 2021 ayant pour objectif d'assurer une collecte de données structurées dans le cadre du nouveau projet d'atlas "Oiseaux De France" (ODF) visant à actualiser l'état des connaissances de l'avifaune française en période de nidification.

Ce protocole vient apporter une réponse aux exigences de Rapportage, dans le cadre de l'application de la Directive Oiseaux de la Commission Européenne. Au titre de l'article 12 de la Directive Oiseaux, les États membres doivent produire, tous les six ans, un rapport sur l'état des populations d'oiseaux. En sa qualité de membre du groupe de coordination nationale, la LPO a en charge la mise à disposition de jeux de données, ainsi que la validation des fiches espèces renseignant la distribution et les suivis d'abondances.

L'origine de ce protocole a été suscitée suite à des travaux précédents menés au cours de mon stage de Master 2 qui portaient sur les données collectées par le protocole EPOC durant la période 2017-2019 afin d'analyser les biais liés à la collecte semi-structurée du protocole (Nabias 2020). Ces travaux avaient montré que, telles quelles, les données du protocole EPOC ne permettaient pas d'estimer correctement les tailles de populations d'oiseaux communs, en cause: (i) l'absence de réplicas temporels empêchant d'estimer correctement le processus d'observation; et (ii) une non-représentativité des habitats prospectés liée aux comportements des observateurs et favorisant certains types de milieux (notamment une surprospection des zones urbaines et une sous-prospection des zones de cultures intensives, par rapport à un tirage aléatoire sur l'ensemble du territoire).

Dans cette optique, la LPO a mis en place un plan d'échantillonnage basé sur une grille systématique ainsi qu'un système de complétion afin de maximiser le nombre de localités prospectées sur la période d'atlas. Sur la période 2021-2023, la LPO tire aléatoirement dix grilles "atlas" (10x10km) pour chaque département correspondant à des Primary Sampling Units (PSU), chaque PSU étant découpée en 25 Secondary Sampling Units (SSU) de 2x2km. Un second tirage aléatoire est réalisé parmi ces SSU, afin de sélectionner 10 localités: cinq considérées comme "prioritaires" et cinq considérées comme "réserve". Il est demandé aux observateurs d'échantillonner les cinq SSU prioritaires et de recourir aux SSU de "réserve" dans le cas où les localités "prioritaires" sont inaccessibles (en fonction du terrain, ou de la présence de domaines privés). Les localités à prospecter correspondent aux centroïdes des SSU tirés aléatoirement (Figure 18).

Les observateurs doivent réaliser trois passages au cours de la saison de reproduction (de mars à juin, voire juillet dans les zones de haute montagne). Chaque passage correspond à la réalisation consécutive de trois EPOC sur le centroïde des SSU sélectionnés. Après complétion des SSU, *i.e.* réalisation de 3x3 listes de 5 minutes au cours de la période de reproduction sur cinq centroïdes différents, les PSU sont considérés comme valides et retirés de la sélection aléatoire pour l'année suivante, afin d'augmenter le nombre de sites prospectés. Au vu de la contrainte liée à un protocole plus standardisé, le protocole EPOC-ODF a reçu une participation moindre en termes de quantité de données collectées (Tableau 1).

Tableau 1 : Répartition du nombre de listes complètes (en milliers) réalisées par les dispositifs EPOC et EPOC-ODF entre 2017 et 2023. Dans le cadre de l'EPOC-ODF, le nombre de sites prospectés est indiqué entre parenthèses

	2017	2018	2019	2020	2021	2022	2023
<b>EPOC</b>	4.4	8.3	12.5	18.1	30.4	33.3	31.1
<b>EPOC-ODF</b>	/	/	/	/	9.6 (1.4)	11.1 (1.4)	7.2 (0.8)

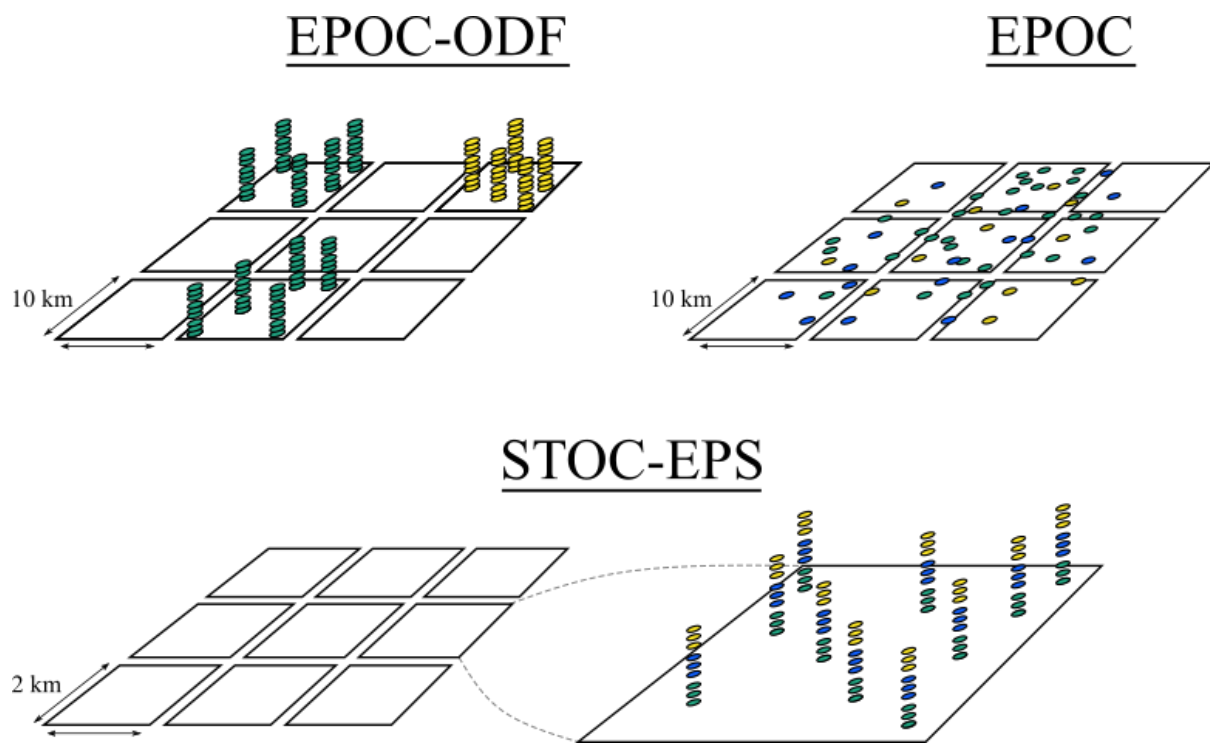


Figure 18 : Représentation schématique des trois jeux de données. Chaque point représentant une instance de collecte de données (pour les trois protocoles, cela correspond à la réalisation d'un point d'écoute de cinq minutes). Les couleurs correspondent à différentes années.

## II.B Méthodes d'analyses

### II.B.i Hierarchical Distance Sampling : L'estimation d'abondance à partir de données de distance

L'inférence des abondances, *i.e.* la détermination des effets des variables environnementales sur leur distribution, peut être décrite à l'aide de modèles espace-état (state-space models ; [Borchers, Buckland & Zucchini, 2002](#); Figure 19). Ces modèles permettent une représentation de la distribution spatiale de l'abondance tout en distinguant les incertitudes liées aux processus écologiques (variabilité du nombre d'individus causée par des facteurs biotiques et abiotiques) des incertitudes liées aux processus d'observation, *e.g.* détection imparfaite ou échantillonnage ([Figure 14](#); [McClintock & Thomas, 2020](#)).

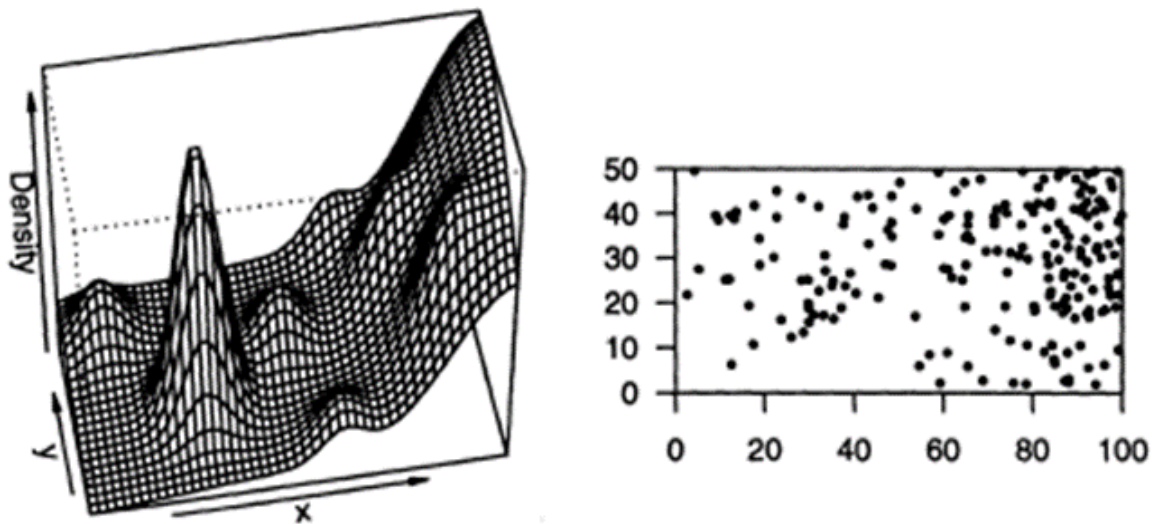


Figure 19 : Représentation d'un modèle spatial d'état selon la position des individus. Les zones de forte densité (concentration d'individus à droite) peuvent correspondre à des habitats préférentiels. Tiré de la figure 3.3 de [Borchers, Buckland & Zucchini, 2002](#).

Le Distance Sampling (DS : [Buckland et al., 1993](#)) est une des méthodes développées afin de prendre en compte des individus potentiellement manqués lors de la visite des observateurs. Le DS a pour objectif d'estimer la densité (voir Boxes 1 et 2) d'individus, en se

basant sur (i) les distances d'observation (observateurs – individus)<sup>19</sup> collectées lors de l'échantillonnage (pouvant être réalisé selon des transects linéaires ou ponctuels ; Figure 20) et (ii) le calibrage d'une fonction de détection  $f(x)$  (cf. Figure 21).

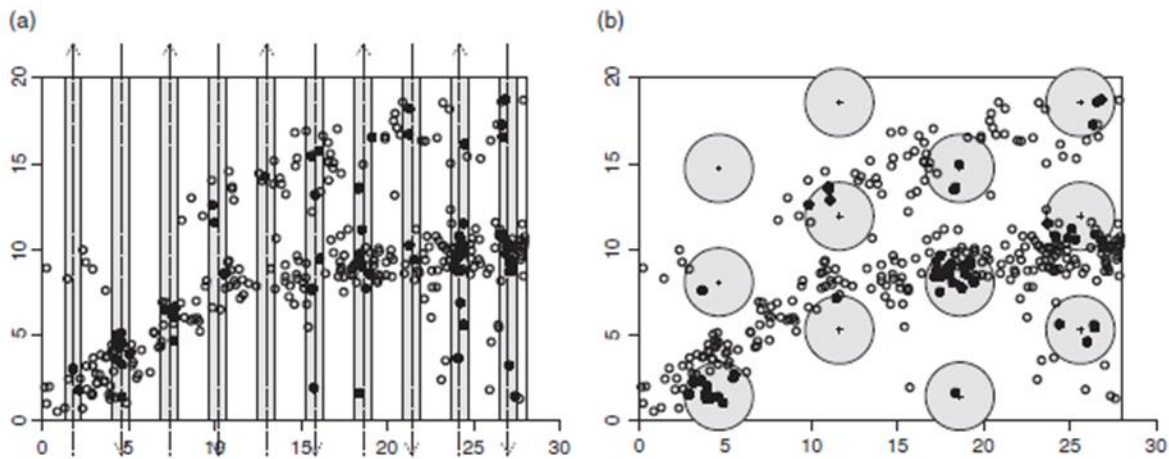


Figure 20 : Représentation schématique de différents types d'échantillonnage (a) linéaire et (b) ponctuel. Les zones grises correspondent aux surfaces prospectées par les observateurs. Les individus détectés et non détectés sont respectivement illustrés par des points noirs et blancs. Tiré de la figure 3.3 de [McClintock & Thomas, 2020](#).

Le DS repose sur quatre hypothèses fortes assurant une inférence non biaisée du nombre d'individus ([Buckland et al., 1993](#)) :

- (i) L'échantillonnage est représentatif<sup>20</sup> de l'environnement prospecté;
- (ii)  $f(x = 0) = 1$  ; Les observateurs détectent parfaitement des individus qui se trouvent à distance nulle;
- (iii) Les individus sont détectés à leur position initiale avant tout mouvement (d'attraction ou d'évitement);
- (iv) Les distances d'observation sont mesurées correctement.

<sup>19</sup> Dans le cadre des transects linéaires, la distance d'observation est calculée à partir du transect et non de la position de l'observateur.

<sup>20</sup> Le terme représentatif sous-entend que les habitats composants la zone d'étude sont échantillonnés équitablement.

Les données de distance collectées sont utilisées pour calibrer une fonction de détection  $f(x)$  décrivant la diminution de la probabilité de détecter un individu selon sa distance de détection (Figure 21A). Cette fonction doit respecter des critères de formes attestant d'une collecte d'information correcte et de la robustesse des estimations, notamment (i) la présence d'un plateau et d'une épaule (asymptote) aux courtes distances et (ii) une décroissance monotone, en relation aux limites physiques des observateurs et de l'atténuation du signal, *e.g.* absorption du signal sonore dans l'environnement (Naguib & Wiley, 2001). La fonction de détection se construit à partir d'une fonction clé '*Key function*' (Buckland et al., 1993) lui permettant de décrire différents profils de décroissance de détectabilité (Figure 21B). Il est possible d'inclure l'effet de variables influençant le profil de détectabilité au travers du paramètre d'échelle ( $\sigma$ ; Marques & Buckland, 2003). Ces variables peuvent être liées à l'effort de collecte ou à des caractéristiques environnementales (La Morgia et al., 2015), comme le type d'habitat (Figure 21C).

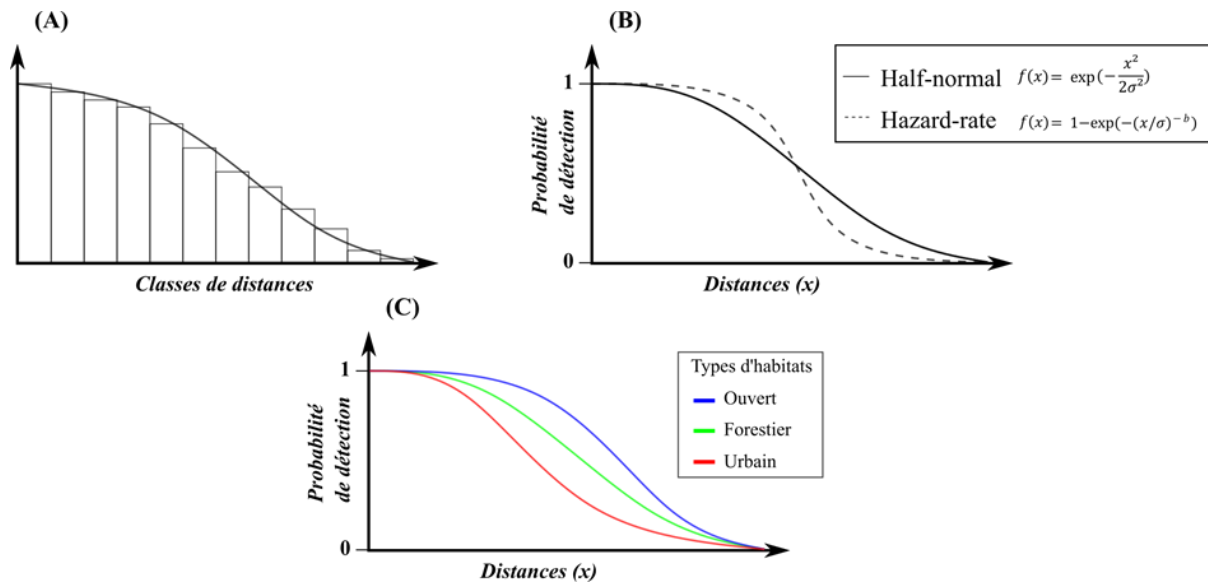


Figure 21 : Représentation de la fonction de détection. Elle se détermine à partir des données de distance d'observation collectées sur le terrain (A). (B) Cette fonction peut s'exprimer selon plusieurs formes, *i.e.* fonctions clés – 'Key functions' (Buckland 1993), afin de décrire différents profils de détection à partir des données collectées. (C) Il est possible de faire varier la fonction de détection selon les conditions environnementales et d'efforts lors de la collecte, ici le type d'habitats, au travers du paramètre ( $\sigma$ ). Dans le cas illustré, à distance égale, les individus ont une probabilité plus importante d'être détectés dans des milieux ouverts que dans des milieux urbains.

Tel quel, le DS ne permet pas d'inférer sur de nouvelles conditions environnementales, hormis au travers d'échantillonnages stratifiés (voir Boxe 1). Au début des années 2000, le DS s'est scindé en deux approches distinctes (A) le DS conditionnel au processus d'observation : modélisation en deux étapes 'two-stage model' (Buckland et al., 2004; Miller et al., 2013) et (B) le Distance Sampling Hiérarchique (HDS ; Royle, Dawson & Bates, 2004). L'approche (A), correspondant à du DS avec modélisation de densité (Distance Sampling – Density Surface Modelling – DS-DSM ; Hedley & Buckland, 2004; Miller et al., 2013), repose sur une modélisation en deux étapes, la première consistant à estimer les densités locales<sup>21</sup>, au travers de la fonction de détection. Ces densités locales sont ensuite utilisées comme variable de réponse dans un modèle additif généralisé (GAM ; Wood, 2017) permettant de prédire sur de nouvelles conditions environnementales (Miller et al., 2013). L'approche (B) repose sur un modèle hiérarchique prenant en compte simultanément (Royle,

<sup>21</sup> Au niveau des sites échantillonnés

Dawson & Bates, 2004; Kéry & Royle, 2015) le processus d'observation, *i.e.* calibration de la fonction de détection, et le processus écologique, *i.e.* variation spatiale de l'abondance (Figure 14). Malgré le fait que l'approche (A) permette l'évaluation distincte des deux processus, facilitant ainsi le diagnostic des modèles (Miller et al., 2013), elle possède un inconvénient majeur sur la non propagation de l'erreur entre les deux processus<sup>22</sup>. Dans le cadre de nos travaux, nous avons utilisé ces deux approches, mais nous nous sommes ensuite focalisés sur l'approche (B) pour sa flexibilité (Kéry & Royle, 2015).

Dans le cadre des HDS, cette flexibilité permet (i) l'ajout de données acquises par différents protocoles d'échantillonnage de natures différentes (Pavanato Julião, 2021; Kéry et al., 2024) et (ii) l'intégration de dynamiques de populations, relâchant ainsi le postulat de population close au travers de l'émigration temporaire, *i.e.* la probabilité qu'un individu soit disponible sur le site de collecte lors de la visite de l'observateur (Chandler, Royle & King, 2011; Kéry & Royle, 2015). L'estimation de la probabilité de disponibilité ( $\phi$ , voir Boxe 2) nécessitant la réalisation de répliques temporels (Chandler, Royle & King, 2011) permet de limiter les fausses absences (erreurs de type II) lors de l'inférence (Tyre et al., 2003) relâchant ainsi l'hypothèse de détection parfaite à distance nulle  $f(x = 0) = 1$  du DS. La prise en compte de l'émigration temporaire représente un niveau de complexité supplémentaire lors de la construction des modèles (Boxe 2) nécessitant l'estimation d'un plus grand nombre de paramètres (Barker et al., 2018). Lors de la réalisation de la liste, les observateurs peuvent aussi renseigner des faux-positifs, *i.e.* la détection d'individus réellement absents (erreur de type I), due soit à une erreur d'identification (Miller et al., 2012; Strickfaden et al., 2020), soit à du double comptage. Tandis que la formation, en amont, d'observateurs permet de réduire les problèmes d'erreur d'identification (Fitzpatrick et al., 2009; Johnston, Matechou & Dennis, 2022), la réalisation de points d'écoute de courte durée

---

<sup>22</sup> A noter que Bravington, Miller & Hedley, 2021 ont présenté une méthode permettant d'ajuster l'incertitude par propagation *ad hoc*

(5 à 10 minutes ; Fuller & Langslow, 1984) permet de réduire les risques de double comptage. De plus, il est possible d'estimer les taux de faux-positifs directement lors de la modélisation (Chambert, Miller & Nichols, 2015).

Cependant, les modèles de DS reposent sur l'hypothèse d'une distribution homogène des individus autour des sites de collecte. Selon la distribution des points de collecte, cela n'est pas forcément respecté. Dans ces cas, il serait nécessaire de prendre en compte l'hétérogénéité des habitats au sein des sites de collecte (Kéry & Royle, 2015; Mizel, Schmidt & Lindberg, 2018) ce qui demanderait une grande finesse dans la résolution des variables d'habitats utilisées.

### Boxe 1: Approches d'estimations basées sur l'échantillonnage et les modèles – Design-based et Model-based

Les premières inférences d'abondances réalisées par le Distance Sampling avaient recours à des inférences basées sur l'échantillonnage 'design-based' nécessitant une stratification représentative de la zone d'étude (Buckland et al., 1993; McClintock & Thomas, 2020). Dans cette approche, l'incertitude liée à l'inférence est 'à la charge' de l'échantillonnage devant prendre en compte la géographie de la zone d'étude et ses diverses conditions environnementales, mais aussi le comportement des animaux et la gestion de la collecte des données par les observateurs (Buckland et al., 1993). L'inférence revient alors à appliquer une règle de 3 et sommer l'ensemble des densités estimées selon les différentes strates par rapport à la surface totale de la zone d'étude :

$$\widehat{D}_v = \frac{E(n)_v}{a_v * P_{a_v}}$$

Et

$$\widehat{D} = \frac{\sum_v A_v * \widehat{D}_v}{A}$$

$v = 1, \dots, V$  : strates de l'échantillonnage

$E(n)_v$  : Nombre d'individu attendus sur  $v$

$a_v$  : Surface prospectée au sein de  $v$

$P_{a_v}$  : Probabilité de détection estimée sur  $v$

$\widehat{D}_v$  : Densité estimée sur  $v$

$A_v$  : Aire totale de  $v$

$A$  : Aire total de la zone d'étude

$\widehat{D}$  : Densité totale estimée sur la zone d'étude

Dans le cas de l'inférence via modélisation, le modèle a pour objectif de décrire une hypothèse de distribution spatiale de l'abondance (Miller et al., 2013) via l'estimation de différents paramètres reliant la quantité d'individus observée aux habitats prospectées (voir Boxe 2). Le recours à des covariables environnementales adéquates permet un apport d'information permettant de diminuer l'incertitude lors de l'inférence à défaut d'une dépendance à l'hypothèse générée par le modèle (McClintock & Thomas, 2020).

Ces deux approches diffèrent par leur manière de représenter l'état des populations. L'inférence par échantillonnage considère que la population est stable durant l'échantillonnage, tandis que l'inférence par modélisation considère que l'échantillonnage est stable, mais que la population observée (et détectée) et une réalisation aléatoire du processus étudié (Dumelle et al., 2022; Aubry & Francesiaz, 2022).

#### Références boîte

- Aubry P, Francesiaz C. 2022. On comparing design-based estimation versus model-based prediction to assess the abundance of biological populations. *Ecological Indicators* 144:109394. DOI: 10.1016/j.ecolind.2022.109394.
- Buckland ST, Anderson DR, Burnham KP, Laake JL. 1993. 3. Statistical theory. In: *Distance Sampling*. Dordrecht: Springer Netherlands,. DOI: 10.1007/978-94-011-1574-2.
- Dumelle M, Higham M, Ver Hoef JM, Olsen AR, Madsen L. 2022. A comparison of design-based and model-based approaches for finite population spatial sampling and inference. *Methods in Ecology and Evolution* 13:2018–2029. DOI: 10.1111/2041-210X.13919.
- McClintock BT, Thomas L. 2020. 3. Estimating Abundance or Occupancy from Unmarked Populations. In: Murray DL, Sandercock BK eds. *Population Ecology in Practice*. Hoboken: Wiley-Blackwell,.
- Miller DL, Burt ML, Rexstad EA, Thomas L. 2013. Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution* 4:1001–1010. DOI: 10.1111/2041-210X.12105.

## Boxe 2: Distance Sampling Hiérarchique prenant en compte l'émigration temporaire

Si l'on considère un échantillonnage prospectant  $n$  sites ( $i = 1, \dots, n$ ) visités  $k$  fois ( $j=1, \dots, k$ ) où les observateurs ont collectés des distances d'observations ( $x_{i,j}$ ) au cours d'observations ( $y_{i,j}$ ). On peut alors décrire le modèle de distance sampling hiérarchique décrivant la distribution spatiale de l'abondance de la population globale  $M_i$  (ou 'superpopulation') composé de plusieurs sous-ensemble de populations  $N_{i,j}$  d'individus disponibles lors des visites  $j$  (Chandler, Royle & King, 2011; Kéry & Royle, 2015). On peut dès lors estimer la superpopulation  $M_i$  par une loi de Poisson de paramètres  $\lambda_i$  décrite par des covariables spécifiques aux sites ( $X_i$ ).

$$M_i \sim \text{Poisson}(\lambda_i)$$

$$\log(\lambda_i) = \beta_0 + \beta * X_i$$

Au cours de chaque visite  $j$ , les sous-ensembles  $N_{i,j}$  de la superpopulation  $M_i$  peuvent être décrit par le tirage binomial suivant de probabilité ( $\varphi_{i,j}$ ) pouvant varier selon les covariables propres aux sites et/ou aux visites<sup>23</sup> ( $U_j$ ).

$$N_{i,j} \sim \text{Binomial}(M_i, \varphi_j)$$

$$\text{logit}(\varphi_j) = \alpha_0 + \alpha * U_{i,j}$$

Cette quantité d'individus disponibles ( $N_{i,j}$ ) va par la suite déterminer le nombre d'individus détectés par les observateurs ( $y_{i,j}$ ) décrit par une distribution multinomiale construite à partir des probabilités de détections relatives regroupées par classes de distances d'observations ( $\pi_{i,j}$ ), d'une fonction de détection  $f(x_{i,j})$  de paramètre d'échelle ( $\sigma_i$ ). Le paramètre d'échelle peut varier selon des covariables propres aux sites et/ou visites ( $Z_i$ ).

$$y_{i,j} \sim \text{Multinomial}(N_{i,j}, \pi_{i,j}), \text{ avec } \pi_{i,j} = f(x_{i,j}; \sigma_i)$$

$$\log(\sigma_i) = \gamma_0 + \gamma * Z_i$$

### Références boîte :

- Chandler RB, Royle JA, King DI. 2011. Inference about density and temporary emigration in unmarked populations. *Ecology* 92:1429–1435. DOI: 10.1890/10-2433.1.
- Kéry M, Royle JA. 2015. 9.5. Open HDS models: Temporary emigration. In: *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models*. Amsterdam ; Boston: Academic Press,.

<sup>23</sup> Dans le cadre des suivis de nidification, le jour julien correspond à l'une des variables d'effort pouvant expliquer les variations au sein de la période de reproduction. En effet, on s'attend à ce que les visites réalisés en début et fin de phénologie de reproduction décomptent moins d'individus que celles réalisées au milieu de la période (Moussus et al., 2009).

## II.B.ii Environnement et modèles d'études

### II.B.ii.a *Espèces cibles*

En France métropolitaine, on dénombre plus de 300 espèces d'oiseaux nichant sur son territoire (UICN France et al., 2016). Nos travaux portent principalement sur les espèces communes, espèces à large aire de distribution et/ou localement abondantes (Rabinowitz, 1981) dont les estimations de taille de population sont jugées comme étant de qualité moyenne (Comolet-Tirman et al., 2022). Le défaut de qualité de ces estimations vient principalement de l'obsolescence<sup>24</sup> ou du manque de données quantitatives acquises lors du dernier atlas des oiseaux communs (Comolet-Tirman et al., 2015). Cette qualité d'estimations déficiente a plusieurs origines ; (i) une démarche semi-quantitative ne prenant pas en compte le processus de détection et les relations des espèces aux habitats (Roché, Muller & Sibley, 2013) ; (ii) leurs abondances et/ou aires de distribution ne permettant pas un suivi exhaustif (Quaintenne et al., 2020) ; (iii) leur statut de conservation, principalement de préoccupation mineure (LC ; UICN France et al., 2016).

---

<sup>24</sup> Comolet-Tirman et al., 2015 fait l'état d'espèce dont l'état de connaissance est relativement bon, mais ayant des données quantitatives potentiellement incomplètes ou périmées.

*II.B.ii.b Spécificité du territoire français métropolitain*

De par sa localisation géographique, le territoire métropolitain se trouve est traversé par deux voies de migration principales reliant l'Eurasie de l'Ouest à l'Afrique ; le couloir Est-Atlantique et le couloir Mer Noire-Mer Méditerranée, reliant l'Eurasie de l'Ouest à l'Afrique (Figure 22A ; Davidson & Stroud, 2006).

Au niveau climatique, la France métropolitaine se caractérise par l'alternance d'un climat océanique sur la façade ouest (faible amplitude thermique au cours de l'année et des précipitations abondantes), d'un climat continental dans la partie est (alternance d'été chaud et d'hiver froid avec un faible taux de précipitations) et enfin d'un climat méditerranéen dans le sud-est (étés chauds et hivers doux, faibles précipitations), sans oublier la présence de massif montagneux important (Figure 22B, Joly et al., 2010). En plus de cette diversité climatique, le territoire métropolitain regroupe différents types d'habitats, bien que majoritairement forestiers et agricoles (Figure 22C). La combinaison des conditions climatiques, la pluralité des types d'habitats et la position géographique du territoire métropolitain permettent à celui-ci d'accueillir une forte diversité d'espèces tout au long de l'année (Jourde, 2015), à la recherche de ressources alimentaires (Heus, 2013). Cela va des espèces présentes en hiver sur la façade ouest, attirées par les faibles amplitudes thermiques, aux espèces méditerranéennes en limite nord de leur aire de répartition.

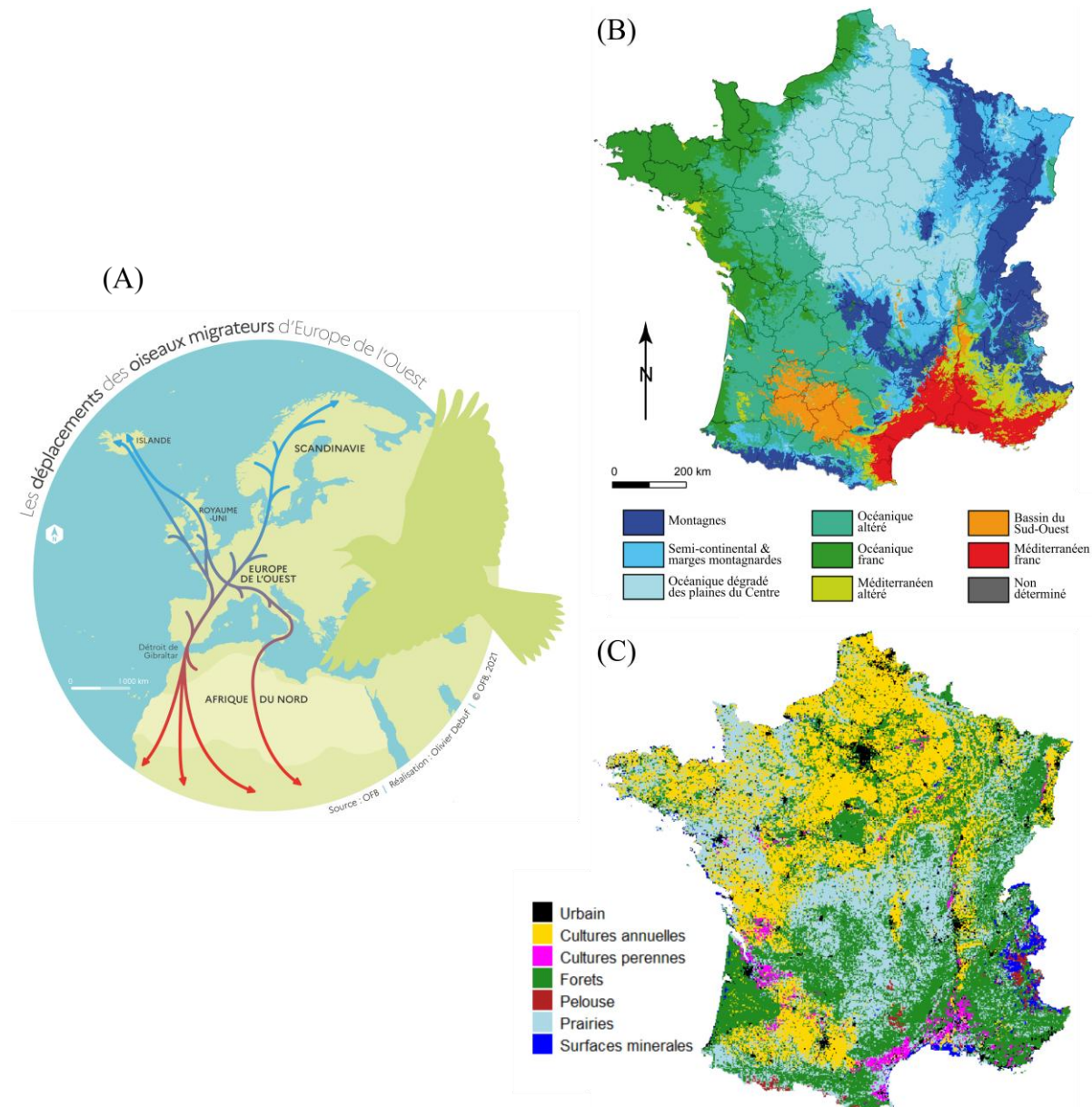


Figure 22 : Figures représentant le contexte géographique et la diversité climatique et de types d'habitats en France métropolitaine. En (A) une représentation des couloirs migratoires entre l'ouest de l'Eurasie et l'Afrique, tiré de [Marcel et al., 2021](#), illustration d'Olivier Debuf. En (B) la carte de répartition des différents types de climats métropolitains, tirée de [Joly et al., 2010](#). En (C) la répartition des différents types d'habitats agrégés à partir de la classification OSO 2020 ([Thierion, Vincent & Valero, 2022](#)).

II.B.ii.c Réduction des données environnementales et quantification  
de l'extrapolation

Cependant, cette grande variété d'habitats peut engendrer un souci lors du calibrage du modèle, celui de la multicolinéarité, *i.e.* la dépendance linéaire d'une ou plusieurs variables, qui peut biaiser les estimations des coefficients du modèle (Mela & Kopalle, 2002). L'objectif principal de nos travaux étant d'extrapoler l'abondance d'individus sur des conditions environnementales non prospectées (Elith & Leathwick, 2009), il était nécessaire d'intégrer un maximum d'informations environnementales tout en limitant les biais liés à la multicolinéarité. Dans ce but, nous avons recouru à la méthode d'Analyse par Composante Principale (ACP) afin de réduire le nombre de variables décrivant la distribution<sup>25</sup> tout en conservant un maximum d'informations (Jolliffe & Cadima, 2016). L'ACP se traduit par la formation d'axes, ou plans, appelés composantes principales, synthétisant au mieux le nuage de points à  $U$  dimensions composé de  $U$  variables décrivant la composition des habitats (Pearson, 1901). Cette réduction du nombre de variables a pour objectif de favoriser la parcimonie du modèle (Coelho, Diniz-Filho & Rangel, 2019) et d'éviter son sur-calibrage qui limite sa transférabilité - *l'extrapolation* - à de nouvelles conditions environnementales (Randin et al., 2006). Cette approche possède une lacune majeure: la description de gradients environnementaux selon des axes linéaires<sup>26</sup> (Jolliffe & Cadima, 2016) ne permet pas d'inclure et d'interpréter facilement les relations quadratiques décrivant des optimums environnementaux (Hutchinson, 1957; Hirzel & Le Lay, 2008).

Malgré les efforts mis en place pour assurer la parcimonie du modèle et favoriser son extrapolation, la prédiction d'abondance sur des conditions environnementales non

---

<sup>25</sup> Dans notre cas, la variabilité de la distribution correspond à la description des sites de collectes, et sites de prédiction, selon la proportion des différents types d'habitats qui les composent.

<sup>26</sup> Dans nos travaux, on établira des gradients environnementaux décrivant des continuums d'habitats selon des gradients d'habitats principalement forestiers à des habitats principalement agricoles.

échantillonnées peut entraîner des estimations biaisées des tailles de populations (Conn, Johnson & Boyeng, 2015). Une partie de nos travaux a donc pour objectif d'estimer et évaluer cette extrapolation. Pour cela, nous avons comparé l'espace environnemental décrit par l'échantillonnage à celui décrivant la zone de prédiction, qui est dans notre cas l'ensemble du territoire français métropolitain (Figure 23). Mesgaran, Cousens & Webber, 2014 décrivent trois cas de figures d'extrapolation lors d'une prédiction sur un nouvel espace environnemental:

- Une extrapolation analogue, où l'espace environnemental ciblé par la prédiction est inclus au sein de l'espace environnemental décrit par l'échantillonnage (points bleus vides – Figure 23).
- L'extrapolation univariée (Novelty type – NT1), où les conditions environnementales de la zone prédite ne sont pas incluses dans un des axes décrivant l'espace environnemental illustré par l'échantillonnage (points noirs – Figure 23).
- L'extrapolation combinatoire (NT2), où l'espace des conditions environnementales de la zone prédite est inclus au sein des étendues des variables environnementales, mais dont la combinaison des variables n'est pas comprise dans l'espace environnemental décrit par l'échantillonnage (points bleus pleins – Figure 23).

Dans nos travaux, nous avons recouru à l'évaluation de l'extrapolation dans le traitement *ad hoc* des estimations aberrantes prédites par les modèles HDS.

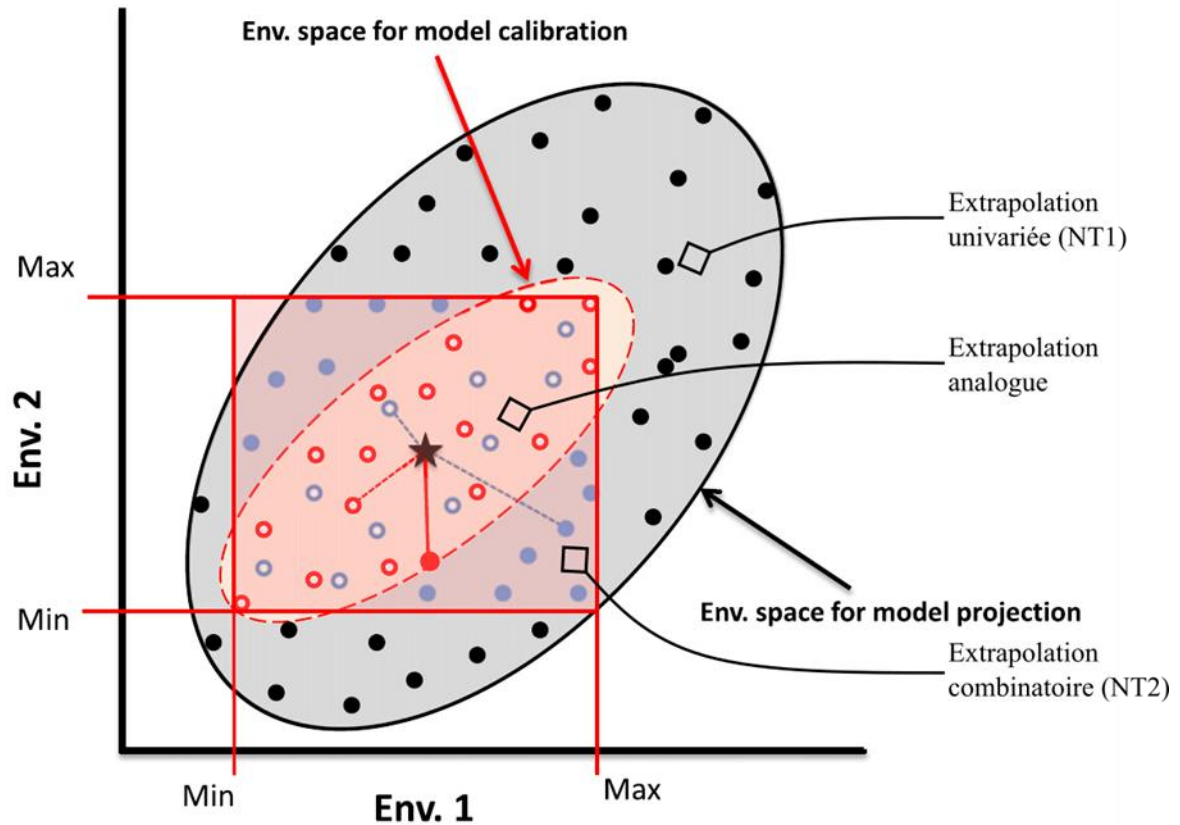


Figure 23 : Représentation schématique d'un espace environnemental décrit par deux variables (Env 1 et Env 2). Les points rouges représentent les données environnementales collectées par l'échantillonnage. L'ovale rouge correspond à l'espace environnemental utilisé lors de la calibration du modèle. Le rectangle rouge représente l'étendue des conditions environnementales collectées par le protocole d'échantillonnage. L'ovale noir représente l'espace environnemental décrivant les conditions environnementales de la zone de prédiction. L'étoile noire montre le barycentre de l'espace environnemental de l'échantillonnage (ovale rouge) utilisé pour calculer les distances entre celui-ci et les positionnement dans l'espace environnemental; pour plus d'informations voir (Mesgaran, Cousens & Webber, 2014). Figure modifiée de la figure 2 de Mesgaran et al., 2014.

II.B.iii Approche d'estimation des paramètres

Dans le cadre de nos inférences, nous avons recouru à la fois aux approches fréquentiste et bayésienne afin d'estimer les paramètres liés aux processus écologiques et d'observation (Figure 14). L'approche fréquentiste repose sur l'estimation de la combinaison de paramètres la plus vraisemblable permettant de reproduire les données observées (Casella & Berger, 2001). Elle permet d'estimer des intervalles de confiance contenant ces paramètres, quelle que soit leur distribution réelle<sup>27</sup> (Efron, 2012). L'approche bayésienne repose, elle, sur (i) des connaissances à priori<sup>28</sup> du processus étudié (ii) la vraisemblance de la description du processus (dans notre cas, la formulation des modèles hiérarchiques) et (iii) les données collectées par l'échantillonnage constituant une description partielle du processus étudié (Gelman et al., 2013). Elle utilise le principe de Bayes, où l'on considère  $p(\theta)$  la distribution à priori des paramètres régissant le processus étudié ;  $p(y|\theta)$  la distribution des données collectées ( $y$ ) conditionnée par ( $\theta$ ) ; et  $p(\theta|y)$  la distribution à postériori des paramètres (conditionnée par les données) :

$$p(\theta|y) \propto p(\theta)p(y|\theta)$$

L'inférence bayésienne nécessite de générer des valeurs aléatoires pour l'ensemble des paramètres du processus étudié en se basant sur leur distribution à priori (Gelman et al., 2013). La génération de cet aléatoire repose principalement sur des simulations de chaînes de Markov – Markov Chain Monte Carlo ; MCMC (Gilks, Richardson & Spiegelhalter, 1995; Gelman et al., 2013), à l'aide de différents algorithmes d'échantillonnage, comme

<sup>27</sup> La figure 5 des travaux d'Efron de 2012 illustre bien ce propos (Efron, 2012)

<sup>28</sup> Elles prennent la forme de distribution décrivant les attendues théoriques des paramètres étudiés, aussi appelé priors. Ces attendues peuvent être plus au moins renseignés selon des connaissances antérieures acquises – *informative priors* –, ou non informelles – *uninformative priors*. Par exemple, dans le cas d'une étude sur l'aire de distribution du pic vert et de l'alouette des champs, on pourrait *a priori* s'attendre à ce que le pic vert soit plus présent en milieux forestier et l'alouette en milieux agricoles en se basant sur les connaissances écologiques des deux espèces. Dans le cas de nos travaux, nous avons principalement utilisé des prior non informelles.

l'algorithme de Metropolis-Hasting, ou l'échantillonneur de Gibbs (Gelman et al., 2013). Dans la pratique, afin de s'assurer de la robustesse des paramètres estimés, il est recommandé de générer plusieurs chaînes de Markov et d'évaluer leurs convergences communes lors de la phase stationnaire des chaînes (Gelman et al., 2013).

Les chaînes de Markov générées peuvent être plus ou moins longues et se distinguent en deux sections, une phase de chauffe, ou 'burn-in', et une phase stationnaire. Durant la phase de chauffe, il y a une prospection de l'espace paramétrique délimité par la distribution à priori. Au fil des itérations, et de l'actualisation de la distribution probable des paramètres, au regard des données collectées, les valeurs générées vont se recentrer sur la distribution postérieure la plus vraisemblable de générer le processus.

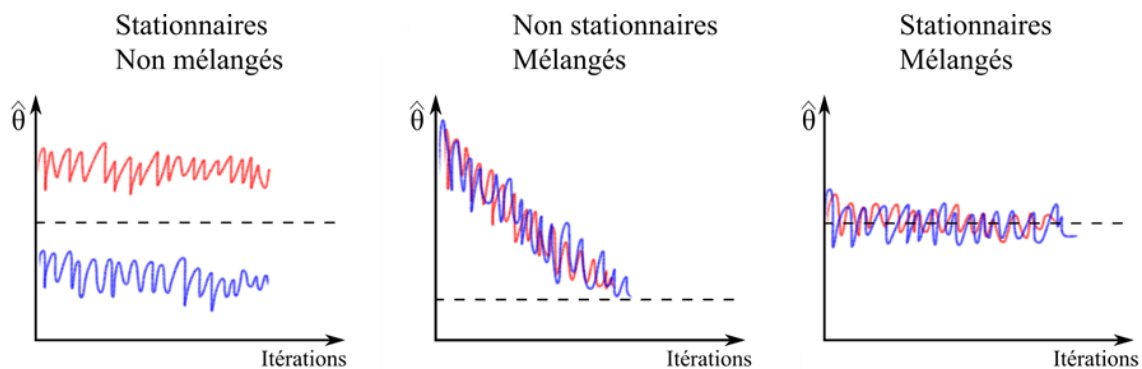


Figure 24 : Représentation schématique de différents cas d'étude pour l'inférence d'un paramètre ( $\hat{\theta}$ ). Après calibration du modèle bayésien, il se peut que les chaînes soient stationnaires, mais non mélangées, décrivant ainsi deux potentielles distributions postérieures pour le paramètre  $\theta$ . Il se peut qu'elles n'aient pas encore atteint la phase stationnaire. La ligne horizontale en pointillé représente la valeur attendue de  $\hat{\theta}$ .

Une inférence bayésienne est considérée comme correcte lorsque les différentes chaînes générées ont convergé, *i.e.* qu'elles décrivent des distributions postérieures du paramètre d'estimation similaires (graphique de droite sur la Figure 24). L'évaluation de la

convergence des chaînes se réalise via le calcul de l'indice Gelman-Rubin<sup>29</sup> ( $\hat{R}$ ; Gelman & Rubin, 1992) correspondant au ratio de la variance inter-chaîne sur la variance intra-chaîne.

L'inférence bayésienne correspond mieux à la démarche analytique de l'écologie de la conservation que l'inférence fréquentiste, car le processus étudié, *e.g.*, l'estimation du nombre d'individus ou l'aire de distribution de l'espèce, prend en compte l'incertitude au regard des données collectées (Ellison, 2004). L'approche fréquentiste est quant à elle plus adéquate pour valider ou invalider des hypothèses davantage propices aux études expérimentales (Tilman et al., 2001; Stephens, Buskirk & del Rio, 2007). La popularité de l'approche fréquentiste peut s'expliquer par la forte demande computationnelle de l'inférence bayésienne pour les algorithmes MCMC (Efron, 2012) et le recours à une connaissance à priori pouvant impacter l'inférence (Aldrich, 1997; Efron, 2012). Dans le cadre des modèles hiérarchiques, la capacité de l'approche bayésienne à explorer un grand nombre de paramètres (Gilks, Richardson & Spiegelhalter, 1995) en fait donc un outil essentiel dans l'inférence de processus écologique via les modèles hiérarchiques (Kéry & Schaub, 2011). De plus, le développement de récentes méthodes comme l'approximation d'inférence bayésienne via Integrated Nested Laplace Approximations (INLA ; Rue, Martino & Chopin, 2009; Van Niekerk et al., 2023) permet de diminuer considérablement le temps de calcul.

---

<sup>29</sup> A noter, que certains travaux récents de Vehtari et al., 2021 proposant une alternative à cet indice d'évaluation de convergence des chaînes

II.B.iv Diagnostic des modèles

Lors du calibrage de modèles à des fins de prédiction, il est important de s'assurer que le modèle puisse prédire correctement le processus étudié dans de nouvelles zones géographiques et/ou périodes temporelles (Tredennick et al., 2021). La validation du modèle, *i.e.* l'évaluation de ses performances prédictives, consiste à évaluer la réponse obtenue par ce modèle à une observation provenant du même jeu de données, ou d'un jeu de données indépendant (Hastie, Tibshirani & Friedman, 2009). Dans le cas d'une évaluation sur le même jeu de données, on procède à un partitionnement des données, en scindant celui-ci entre les données utilisées pour calibrer le modèle '*train*' et les données utilisées pour l'évaluer '*test*' qui, réalisé plusieurs fois, permet une validation croisée '*cross-validation*' (Hastie, Tibshirani & Friedman, 2009). L'évaluation du modèle nécessite une mesure de quantification de la divergence entre la réponse du modèle et une observation indépendante<sup>30</sup>. Typiquement, dans le cadre des modèles de distribution d'espèces nécessitant des données de présence/absence pour déterminer les localités favorables à la présence d'une espèce, on va recourir à des métriques comme le ROC-AUC (*Receiver-Operator-Characteristic - Area Under the Curve* ; Junge & Dettori, 2018) se basant sur la matrice de confusion entre les présences/absences prédites et observées (Peterson, Papeş & Soberón, 2008).

Dans le cadre de mes travaux, j'ai évalué le calibrage des modèles '*goodness-of-fit*' en utilisant la mesure statistique du  $\chi^2$  comparant les mesures observées aux mesures estimées par les modèles<sup>31</sup>. J'ai utilisé une méthode de bootstrap paramétrique afin de recréer un nouveau jeu de données construit à partir des paramètres estimés par le modèle (Borchers, Buckland & Zucchini, 2002; Kéry & Royle, 2015). Ce nouveau jeu de données simulé est

<sup>30</sup> Ici, les termes 'd'observation indépendante' renvoient au fait que les données ne sont utilisées pour calibrer le modèle. Dans le cas des validations croisées, les données d'observations ne sont pas réellement indépendantes, car elles sont régies par le même processus d'observation que les données de calibrage (Figure 14).

<sup>31</sup> Pour chaque site  $i = \{1, \dots, n\}$ ,  $\chi^2 = \sum_1^n \frac{(\text{Observée}_i - \text{Estimée}_i)^2}{\text{Estimée}_i}$

ensuite utilisé pour calibrer un nouveau modèle permettant de comparer les valeurs simulées et nouvellement estimées. Ce processus est réalisé plusieurs fois afin d'obtenir une distribution de la valeur du  $\chi^2$  paramétrique permettant de tester si le  $\chi^2$  initialement calculé est inclus dans la distribution des  $\chi^2$  paramétriques (Kéry & Royle, 2015). Lors de l'inférence bayésienne, en plus de l'estimation de la distribution postérieure des paramètres du modèle, j'ai utilisé ces mêmes paramètres afin de simuler des données attendues. A chaque itération j'ai calculé les deux métriques de divergence, l'une basée sur les données d'observation et les paramètres estimés par le modèle ( $\chi_{obs}^2$ ) et l'autre basée sur les données attendues et les paramètres estimés ( $\chi_{attendue}^2$ ), afin de tester l'hypothèse nulle  $\chi_{attendue}^2 \geq \chi_{obs}^2$  où on prédit que la proportion de fois où  $\chi_{attendue}^2$  est supérieure à  $\chi_{obs}^2$  soit égale à 0.5 dans le cas d'un modèle correctement calibré (Kéry & Royle, 2015).

## II.B.v Sélection des modèles

*'The tendency of the mind, accustomed to work through multiple hypotheses, is to sway to one line of policy or another, according as the balance of evidence shall incline.'*, T.C. Chamberlin, 1890

De par la complexité des processus écologiques, plusieurs hypothèses alternatives sur l'effet de facteurs<sup>32</sup> peuvent être adéquates pour les expliquer (Hilborn & Mangel, 1997; Johnson & Omland, 2004). De plus, l'expansion d'acquisition des données via satellites 'remote sensing' accroît considérablement la quantité de descripteurs d'habitats disponibles (Crowley & Cardille, 2020). L'inférence multimodale permet de prendre en compte les paramètres estimés par plusieurs modèles dans l'optique de produire un résultat consensuel (Johnson & Omland, 2004; Murray et al., 2020). Elle se base sur la théorie de l'information

---

<sup>32</sup> Ces facteurs peuvent être biotiques, dû à l'interaction avec d'autres espèces (comme la compétition), ou abiotiques (lié à l'habitat physique). Dans nos travaux, nous nous focalisons principalement sur les facteurs abiotiques

(Kullback-Leibler information ; [Kullback & Leibler, 1951](#)) permettant d'estimer l'évidence relative des modèles les uns par rapport aux autres (Figure 25A; [Burnham & Anderson, 2001](#); [Burnham, Anderson & Huyvaert, 2011](#)). La quantification de l'évidence des modèles se base sur des mesures prenant en compte la vraisemblance comme le critère d'Akaike (Akaike Information Criterion, AIC [Akaike, 1974](#)) ou BIC (Bayesian Information Criterion ; [Schwarz, 1978](#)) pénalisant plus au moins la complexité (ici, le nombre de paramètres utilisés) du modèle. Cette quantification permet d'ordonner les modèles au regard de leurs évidences (via leur différence dans la figure ;  $\Delta AIC$  Figure 25B).

Dans le cadre des modèles hiérarchiques, composés de plusieurs modèles, tester l'ensemble de combinaisons pouvant expliquer les différents processus impliqués peut rapidement devenir très demandeur en ressource computationnelle ([Morin et al., 2020](#)). L'aspect imbriqué de ces modèles nécessite le recours à des méthodes de sélection, comme celle de la stratégie du second ensemble de modèles candidats '*Second candidat sets models*', ([Morin et al., 2020](#)) visant à réaliser deux étapes de sélection. Une première étape aboutit à l'ensemble des modèles expliquant convenablement ( $\delta_{AIC} \leq 10$ ) les différents processus indépendamment des autres processus (modéliser selon l'hypothèse nulle). Suite à cela, les modèles retenus de chaque processus sont ensuite réutilisés dans une seconde étape de sélection, qui mène à l'ensemble de modèles final composé des modèles établis à partir de la combinaison des modèles sélectionnés ([Bromaghin, McDonald & Amstrup, 2013](#))

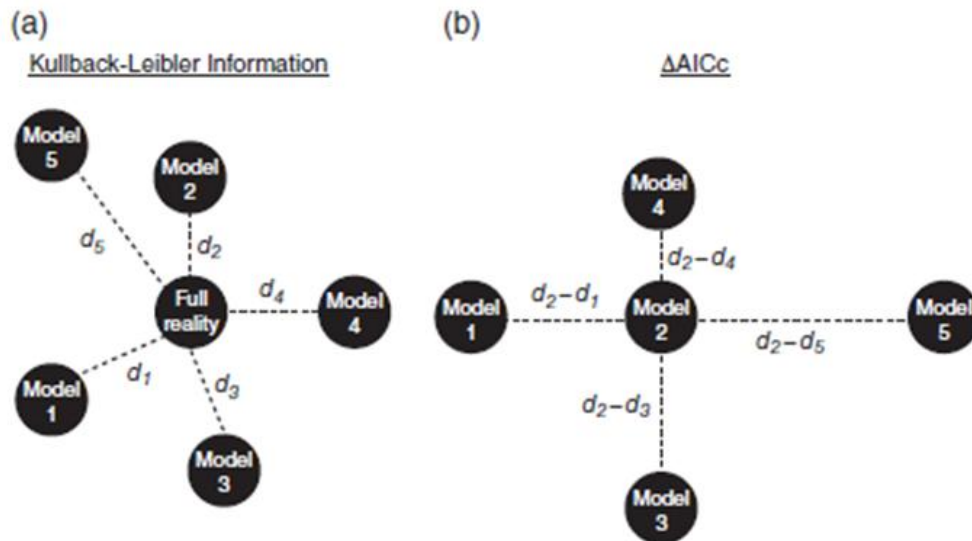


Figure 25 : Représentation schématique de la théorie de l'information, en prenant l'exemple de l'AIC. Les différents modèles (ou hypothèses) sont plus ou moins distants de la 'réalité' attendue (a). Cette 'réalité' étant inconnue, le modèle avec la plus courte distance (AIC le plus faible) de l'ensemble des modèles testés correspond au modèle s'y rapprochant le plus (b). Dans l'inférence multimodale, les paramètres du modèle 2 auront plus de poids que les paramètres estimés par les autres modèles. Figure tirée de Murray et al., 2020.

Malgré l'attrait de cette approche, son objectivité apparente (Chamberlin, 1890) et le besoin d'évidence dans la mise en place de mesures de conservation de la biodiversité (Rehme, Powell & Allen, 2011), l'inférence multimodale possède des défauts (Burnham, Anderson & Huyyaert, 2011). De plus, le recours systématique à l'inférence multimodale, associé à l'essor du remote-sensing peut dévier les écologues de la pratique d'établir en amont des hypothèses basées sur leurs connaissances du modèle étudié ou à priori du processus (Gotelli & Ellison, 2012; Murray et al., 2020). L'inférence multiple peut aussi avantager les modèles décrivant au mieux les données, favorisant ainsi les modèles surcalibrés (Murray et al., 2020) et réduisant la transférabilité des réponses calibrées à d'autres conditions environnementales (Randin et al., 2006).



# III Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives

---

## III.A Contexte

Afin de mesurer les avancées dans l'application de la Directive Oiseaux, la Commission Européenne demande aux États membres de réaliser tous les six ans un rapport sur l'état des populations d'oiseaux nicheurs (Article 12 ; Directive 2009/147/CE du Parlement européen et du Conseil du 30 novembre 2009 concernant la conservation des oiseaux sauvages, 2009). En tant que membre du groupe de coordination nationale 'Directive Oiseaux', la LPO a pour charge (i) la mise à disposition de jeux de données permettant le suivi d'abondance et de distribution, ainsi que (ii) la validation de fiches espèces restituant l'ensemble des données transmises par la France à la Commission européenne.

Suite au premier exercice de rapportage réalisé sur la période 2008-2012 résultant dans la réalisation de l'Atlas des Oiseaux nicheurs de France Métropolitaine (AOFM ; Issa & Muller, 2015) et aux difficultés rencontrées lors de la collecte et l'exploitation des données (Comolet-Tirman et al., 2015), la LPO a décidé de mettre en place un dispositif de collecte basé sur les sciences participatives de données semi-structurées (EPOC) et structurées (EPOC-ODF) afin d'assurer la production d'effectifs nicheurs affinés et traçables.

Dans ce premier chapitre, nous avons utilisé uniquement les données d'EPOC-ODF pour l'estimation des effectifs nicheurs des 54 espèces d'oiseaux les mieux échantillonnées en France. Bien que les programmes EPOC et EPOC-ODF soient similaires sur certains aspects de leurs protocoles, à savoir la localisation de positions des individus autour des

## Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives Contexte

observateurs durant un point d'écoute de cinq minutes, les deux programmes diffèrent sur deux aspects (i) le choix de la localité n'est pas imposé pour le protocole EPOC et (ii) le protocole EPOC-ODF requiert la réalisation de réplicas temporels au cours de la période de reproduction. À l'aide des données collectées par EPOC-ODF, nous avons mis en place une approche de modélisation globale à l'ensemble des espèces communes nicheuses basée sur des modèles de Distance Sampling Hiérarchique (HDS ; [Chandler, Royle & King, 2011](#)). Ces modèles permettent d'expliquer la variation spatiale d'abondance, *processus écologique*, en tenant compte des variations liées à la collecte de données, *processus d'observation* (Figure 14). La comparaison des résultats issus du précédent atlas ([Issa & Muller, 2015](#)) après actualisation à l'aide des tendances temporelles issues du STOC-EPS ([Vimont et al., Submitted](#)) aux estimations obtenues par les données EPOC-ODF montre que 65% des 46 espèces retenues<sup>33</sup> possédaient des estimations d'effectifs nicheurs inférieures dans l'atlas précédent. Ces résultats démontrent une potentielle sous-estimation généralisée par l'ancienne approche ([Roché, Muller & Sibley, 2013](#)), à l'exception notable de certaines espèces dont la détectabilité est forte, car leur chant porte loin et qui sont, elles, potentiellement surestimées. Ces résultats montrent l'importance de la prise en compte du processus d'observation pour la surveillance des populations.

### Référence du contexte

Chandler RB, Royle JA, King DI. 2011. Inference about density and temporary emigration in unmarked populations. *Ecology* 92:1429–1435. DOI: 10.1890/10-2433.1.

Comolet-Tirman J, Sibley J-P, Witté I, Cadiou B, Czajkowski M-A, Deceuninck B, Jiguet F, Landry P, Quaintenne G, Roché J-E, Sarasa M, Touroult J. 2015. STATUTS ET TENDANCES DES POPULATIONS D'OISEAUX NICHEURS DE FRANCE Bilan simplifié du premier rapportage national au titre de la Directive Oiseaux. 83.

Directive 2009/147/CE du Parlement européen et du Conseil du 30 novembre 2009 concernant la conservation des oiseaux sauvages (Version codifiée). 2009.

Issa N, Muller Y. 2015. Atlas des oiseaux de France métropolitaine: Nidification et présence hivernale. Paris: DELACHAUX.

Roché J-E, Muller Y, Sibley J-P. 2013. Une méthode simple pour estimer les populations d'oiseaux communs nicheurs en France. *Alauda* 81:241–268.

<sup>33</sup> Parmi l'ensemble des 54 espèces avec des modèles jugés convenables, nous avons décidé de retirer huit espèces qui montraient des signes de divergences importantes si l'on restreint ou non les surfaces prédites aux étendues spatiales où les espèces étaient considérées comme nicheuses en 2012.

III.B Article

## **Reassessment of French breeding bird population sizes using citizen science and accounting for species detectability**

*Publié dans PeerJ ([10.7717/peerj.17889](https://doi.org/10.7717/peerj.17889))*

Jean Nabias<sup>1,2</sup>, Luc Barbaro<sup>2,3</sup>, Benoit Fontaine<sup>2,4</sup>, Jérémy Dupuy<sup>1</sup>, Laurent Couzi<sup>1</sup>, Clément Vallé<sup>2</sup>, Romain Lorrillière<sup>2,5</sup>

1: LPO-BirdLife France, Fonderies Royales, Rochefort Cedex, France

2: CESCO, Muséum National d'Histoire Naturelle, CNRS, Sorbonne-University, Paris, France

3: Dynafor, INRA-INPT, University of Toulouse – Auzeville, France

4: PATRINAT, OFB, Muséum National d'Histoire Naturelle, CNRS – Paris, France

5: Centre de Recherches sur la Biologie des Populations d'Oiseaux (CRBPO), MNHN-CNRS-OFB, Paris, France

Keywords: Bird atlases; Biogeography; Breeding Bird Surveys; Citizen Science; Detectability; Hierarchical Distance Sampling

Higher efficiency in large-scale and long-term biodiversity monitoring can be obtained through the use of Essential Biodiversity Variables, among which species population sizes provide key data for conservation programs. Relevant estimations and assessment of actual population sizes are critical for species conservation, especially in the current context of global biodiversity erosion. However, knowledge on population size varies greatly, depending on species conservation status and ranges. While the most threatened or restricted-range species generally benefit from exhaustive counts and surveys, monitoring common and widespread species population size tends to be neglected or is simply more challenging to achieve. In such a context, citizen science (CS) is a powerful tool for the long-term monitoring of common species through the engagement of various volunteers, permitting data acquisition on the long term and over large spatial scales. Despite this substantially increased sampling effort, detectability issues imply that even common species may remain unnoticed at suitable sites. The use of structured CS schemes, including repeated visits, enables to model the detection process, permitting reliable inferences of population size estimates. Here, we relied on a large French structured CS scheme (EPOC-ODF) comprising 27 156 complete checklists over 3 873 sites collected during the 2021-2023 breeding seasons to estimate the population size of 63 common bird species using Hierarchical Distance Sampling (HDS). These population size estimates were compared to the previous expert-based French breeding bird atlas estimations, which did not account for detectability issues. We found that population size estimates from the former French breeding bird atlas were lower than those estimated using HDS for 65% of species. Such a prevalence of lower estimations is likely due to more conservative estimates inferred from semi-quantitative expert-based assessments used for the previous atlas. We also found that species

with long-range songs such as the Common Cuckoo (*Cuculus canorus*), Eurasian Hoopoe (*Upupa epops*) or the Eurasian Blackbird (*Turdus merula*) had, in contrast, higher estimated population sizes in the previous atlas than in our HDS models. Our study highlights the need to rely on sound statistical methodology to ensure reliable ecological inferences with adequate uncertainty estimation and advocates for a higher reliance on structured CS in support of long-term biodiversity monitoring.

## Introduction

Worldwide bird populations are nowadays caught in the midst of a global, man-driven erosion of biodiversity caused by the synergistic effects of habitat destruction and fragmentation, resource overexploitation, climate change, pollution, pesticide use as well as the global spread of invasive species (Ceballos et al., 2015; Johnson et al., 2017). In Europe, the intensification of agriculture crystallises this phenomenon because the ever-increased use of pesticides and fertilisers has been pointed out as the main driver of current bird population declines (Rigal et al., 2023). Recent researches have pointed out how important species-specific life attributes (e.g., species range and density) and functional traits (e.g., body mass, diet or dispersal) explain long-term bird trends and responses to global changes (Hong et al., 2023; Santini et al., 2023; Germain et al., 2023). Monitoring long-term as well as shorter-term trends have been historically possible through standardised biodiversity - including bird - surveys at a national scale and aggregated at the continental one (Jiguet et al., 2012; Pilotto et al., 2020; Brlík et al., 2021). This monitoring produces comprehensive distribution atlases pointing out current ranges and their historical changes (Keller et al., 2020), as well as long-term population trends (Brlík et al., 2021). Identifying common species declines in the long-term should be accounted for in agricultural and planning policies (Gaston & Fuller, 2008;

Rigal et al., 2023) to help reconciling society's needs with the safeguarding of ordinary biodiversity (Couvét & Ducarme, 2014).

Monitoring the success of the implementation of such policies is possible through the measurement of Essential Biodiversity Variables (Jetz et al., 2019; Navarro et al., 2017; Pereira et al., 2013), including the assessment of species distribution range and population abundance or density (Santini et al., 2023). Global monitoring capabilities of species populations have increased over the past few decades as citizen science (CS) has gained prominence across various domains, particularly in ecology (Kullenberg & Kasperowski, 2016). This participatory approach has provided the public with unprecedented opportunities to contribute to biodiversity monitoring through data collection and indirectly through their engagement in policy evaluation (McKinley et al., 2017). The development of CS can be related to the emergence of multiple online databases (Newman et al., 2012) capable of gathering comprehensive datasets across large spatial scales and multiple taxa such as iNaturalist ([www.inaturalist.org](http://www.inaturalist.org)) and BioloVision ([data.bioloVISION.net](http://data.bioloVISION.net)) or taxa-centred, for instance, eBird ([www.ebird.org](http://www.ebird.org)). The emergence of CS databases has thus resulted in an exponential increase in biodiversity monitoring capabilities, ranging from phenological shifts (Hurlbert & Liang, 2012), species distribution (Suzuki-Ohno et al., 2017; Matutini et al., 2021) and abundance estimation (Callaghan et al., 2021).

Traditionally, species autecology was recensed in atlases providing temporal snapshots of known distribution and population size estimates using discontinuously gathered data collected during a short period (Donald and Fuller, 1998) and allowing long-term changes assessment when repeated over decades (Keller et al., 2020). The first attempt of French avifauna compilation dates back to 1936 (Mayaud et al., 1936) while the first national atlas was published in 1976 (Yeatman et al., 1976) followed by subsequent in 1995

(Yeatman-Berthelot & Jarry, 1995) and 2015 (Issa & Muller, 2015). Each iteration was marked by a significant increase in participating citizen birders, ranging from 500 to 2400 to 10,000. However, while the ultimate aim of bird atlases is to acquire even knowledge across multiple species for a given geographical area (typically national-wide inquiries), it is essential to note that rare and threatened species benefit from more in-depth population monitoring than more frequent and abundant ones (Ingram et al., 2021). Such exhaustive knowledge of rare species populations is due to interest risen by high extinction risk (IUCN, 2001), the need for recovery plans (Farrier et al., 2007) and narrow distributions allowing exhaustive counts (Quaintenne et al., 2020). In contrast, common species (being altogether abundant, widespread; Rabinowitz, 1981) are considered Least Concern (LC). They consequently tend to receive lesser attention than rarer species (Neeson et al., 2018) despite being key components of global avian biomass changes (Gaston & Fuller, 2008; Inger et al., 2015; Whelan et al., 2015; Rigal et al., 2023).

The principal cause of low quantitative coverage of common birds can be tied to the lack of specific funding stemming from a conservation prioritisation approach (Brooks et al., 2006; Meine et al., 2006) and the trade-off between data quality and data acquisition over large spatial scales (Devictor et al., 2010; Kamp et al., 2016). In France, this incomplete knowledge translates into a significant drop in quality for population size estimates of most common bird species, where 60% of breeding species population estimates are qualified as medium (*i.e.* state of knowledge of species abundance considered more or less satisfying, but semi-quantitative data are either lacking or outdated, see Comolet-Tirman et al., 2015) while for rare and localised species (24% of species), estimates are considered as highly reliable (Comolet-Tirman et al., 2022). To account for this discrepancy, a semi-quantitative estimation method was used for the last atlas survey, with population sizes simply inferred

from the average number of breeding pairs measured over 10x10km grid cells using abundance classes of 1-9, 10-99, 100-999, 1000-6666 (the last upper limit could vary according to knowledge about species densities) multiplied by the number of grid cells known for nesting. Population sizes estimated using this method correspond to lower and upper limits. Lower limits were obtained by calculating the geometric mean of the abundance classes weighted by the number of 10x10 grid cells belonging to these abundance classes while upper limits were obtained by using the arithmetic means (see Roché et al., 2013 method referred further as ArGeom).

However, similarly to other studies (Kellner & Swihart, 2014), this approach fails to account for species detectability  $p$  defined as the probability of detecting at least one individual of a given species in a particular sampling effort, given that individuals of that species are present in the area of interest during the sampling session (Boulinier et al., 1998). Numerous studies have previously shown that  $p$  varies with time of day and season (Skirvin, 1981), observers (Quinn et al., 2011) and year-specific factors (Kéry & Schmid, 2004). Omitting species detectability by assuming perfect or constant  $p$  across sampling schemes, observers and habitat types can lead to biased inferences (Nichols et al., 2009; Kéry, 2011) and affect the estimation of long-term trends due to its unaddressed variation (Schmidt et al., 2013; Sanz-Pérez et al., 2020).

Here, we propose an estimation method enabling a more robust approach of population size estimations. We provide associated uncertainty intervals built upon a revised structured sampling scheme, ensuring data traceability and allowing inferences in the spatial variation of species abundance by formally including the detection process within the modelling framework.

In this study, we aimed at testing whether applying this modelling framework on an unprecedentedly large citizen-based dataset collected over France would (i) provide a new quantitative evaluation of French breeding bird populations and (ii) allow a comparison of population sizes inferred through Hierarchical Distance sampling from those inferred using the previous atlas methodology ArGeom across a large part of French avifauna. In particular, we expect that quantifying the influence of species detectability would allow more relevant ecological inferences (e.g., including environmental and sampling effort covariates to the models) to approach closer to a realistic estimation of breeding bird population size at a national level than previously used methodologies.

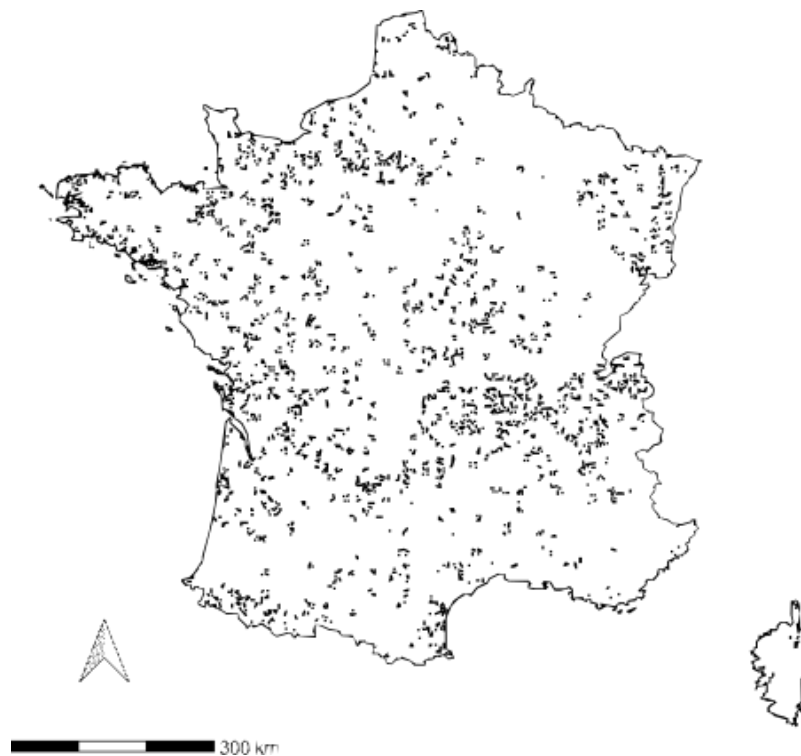
## Materials & Methods

### *Sampling protocol*

EPOC-ODF (French structured estimation of breeding bird population size) is a French CS monitoring scheme based upon 5-minute point counts, where observers are tasked to point locations of recorded individuals, either through visual or auditory detection. Birders can register their field observations directly using the NaturaList smartphone application or transcript later on the data portal Faune-France ([www.faune-france.org](http://www.faune-france.org)). The survey locations corresponded to the centroids of a 2x2km grid, selected from a random sampling. Each location has to be visited three times during the breeding season, from March to June, each consisting of three successive 5-minute point counts, to limit chances of duplicated counts while being less demanding in observation effort (Fuller & Langslow, 1984). After completion, *i.e.* nine visits during a breeding season, surveyed sites are removed from the sampling pool for the subsequent year, to maximise the number of sites surveyed. See Appendix S1 for more details about the sampling design.

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives* *Article*

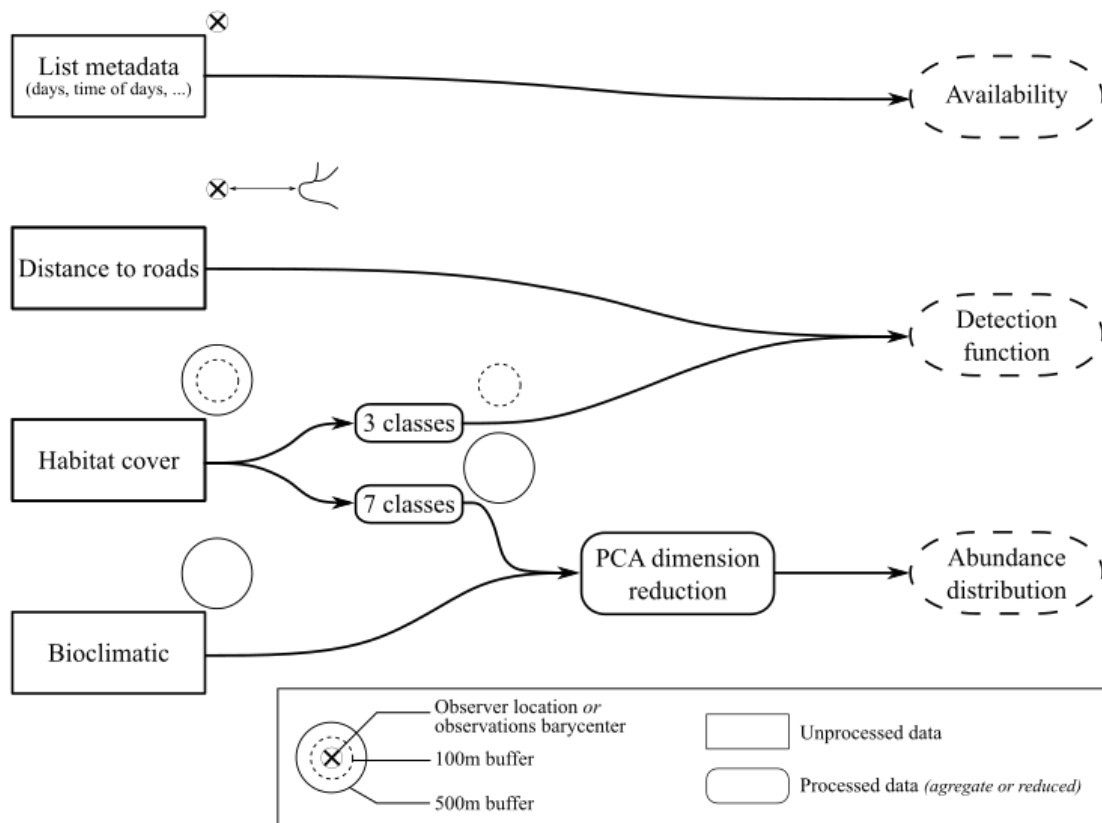
Over the 2021 and 2023 breeding seasons, 276 distinct species were encountered over 27 156 complete checklists collected over 3 873 pre-selected locations (Fig 1) by 520 observers. Sampling effort is monitored through local associations tasked to recruit volunteers. The primary focus of the scheme being the monitoring of common breeding bird species, we decided to constrain the number of species considered viable targets of this scheme to 103 out of the 276 species contacted. We narrowed our study to 63 species out of the initial set of 103, comprising only those recorded at a minimum of 150 distinct locations (3.9% of total locations), to have a sample size allowing to reach model convergence. We also applied a temporal filter that considered both observed activity during the breeding season and expert opinion to capture the breeding phenology of each targeted species and exclude possible early or late migrants from population size estimates (see Appendix S2.1).



**Figure 1 - Spatial distribution of surveyed EPOC-ODF locations (n = 3 505) over 2021-2023 breeding seasons.**

For bioclimatic data, we used 19 variables from WorldClim at 1km resolution (Fick & Hijmans, 2017), on which we applied a Principal Component Analysis (PCA), keeping the first three axes (82.3% of explained variance), to limit multicollinearity through orthogonal transformation (Cruz-Cárdenas et al., 2014).

We used habitat cover data from Theia OSO at 10m resolution (Thierion et al., 2022) and aggregated it according to two different scales: (1) a seven-class corresponding to habitat type (Urban, Annual crops, Perennial crops, Pastures, Grasslands, Forests, Water body/Mineral surfaces) and (2) three-class (Open, Forests, and Artificial) in regards to overall effect on detectability (fig 2). Additionally, we conducted PCA dimension reduction on the seven-class aggregation, retaining three of the six PCA (54.71% of explained variance) axes depicting environmental gradients for (i) forest-to-open-field cultures; (ii) open-field cultures-to-pastures and (iii) perennial crop-to-urban Appendix S3 for the workflow pipeline and habitat cover aggregation. Distances to roads were measured from ROUTE 500 (Cote et al., 2021). Environmental covariates were extracted over a 500m buffer radii upon registered observer location (fig 2). These distances were chosen according to mean dispersal distances and home range sizes in common European birds (Paradis et al., 1998). The three-class habitat covers were collected upon 100m circles radii to assess immediate habitat types that could hinder species detection. Whenever the exact location was unavailable, we used the centroid of sightings as a proxy for observer location (Appendix S4). We used environmental data collected from a prediction grid covering France at a resolution of 2x2km for PCA dimension reduction. Outcomes from this initial PCA were used to transform environmental data collected from surveyed locations through PCA projections.



**Figure 2 - Global overview of covariates acquisition, treatments and usage workflow.** Data are retrieved over observers' GPS location or approximated using observations barycenter, when unavailable, over two resolutions, 100m buffer (dotted circles) and 500m buffer radii (solid circles). Distance to roads is determined by measuring the distance between the nearest road to the observer location or observation barycenter. Habitat cover, in percentage, is aggregate over seven and three classes (see Table S3.1). Seven-class habitat cover and bioclimatic are reduced from PCA keeping the first three dimensions for bioclimatic data and three selected for the seven-class habitat cover.

### Modelling framework

We used Hierarchical Distance Sampling (HDS) models to estimate the abundance of the target species while accounting for uncertainty arising from the observation process (Chandler et al., 2011; Kéry & Royle, 2015). We applied a right-side truncation of 5%, removing observation distances above the 95% quantile, for each targeted species to remove extreme distance values for model robustness (Buckland et al., 2001). Then, we divided observation distances into five proportional bin classes based on the maximal observed distance. Models calibration and assessment were done using unmarked 1.2.5 R package

date and the hours of list realisation (as minutes from sunrise), see Table 1.

Distance sampling key functions, depicting detection probabilities fall-off given distance of observation (Buckland et al., 2001), were chosen between half-normal and hazard-rate based on AIC (Akaike Information Criterion, Akaike, 1974), with other states kept constant.

We based our modelling framework on a secondary candidate set strategy (fig 3), where the detection and availability states of our HDS were fit according to the set of the first candidates while others were kept constant (Morin et al., 2020). For the Poisson process underlying abundance distribution, we used a single model consisting of retained covariate PCAs axes (Table 1). See Appendix Table S5.1 for the number of times where each sub-process was included in the final candidate sets.

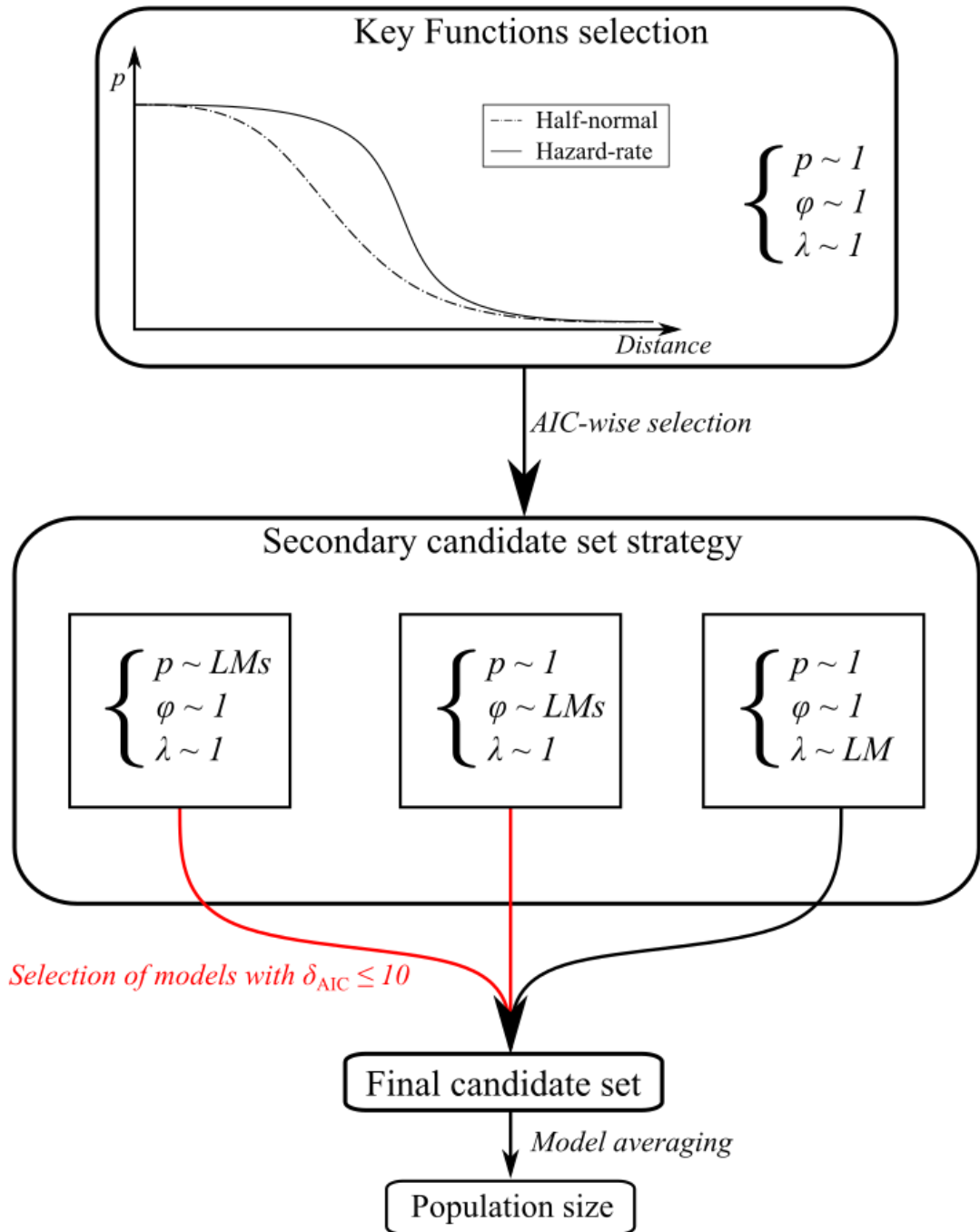


Figure 3 - Methodological framework for population size estimation. At first, a key function is determined AIC-wise between half-normal and hazard-rate with other components of the HDS kept constant. The selected key function is then used during the secondary candidate set strategy (Morin et al., 2020), fitting multiple sub-models (Table 1) for each state separately holding others constant. Sub-models with greatest support,  $AIC \leq 10$  are then selected in a final candidate set consisting of multiple HDS through combinatorial association (Morin et al., 2020). At this stage, we used sub-models estimated coefficients as starting values to help model convergence. Population size estimates are obtained through model averaging of the final candidate set. *LM* : Linear Model

**Table 1 - Ensemble of sub-models tested in the secondary candidate set approach (Morin et al., 2020). (\*) For the hour effort covariate, we used minutes from sunrise estimated from site longitude, latitude and date of list completion.**

States	Sub-models
Detection	~ Distance to roads
	~ Distance to roads + Proportion of artificial lands (100m)
	~ Distance to roads + Proportion of open lands (100m)
	~ Distance to roads + Proportion of forests (100m)
Availability	~ Julian date
	~ Julian date + Julian date <sup>2</sup>
	~ Hour (*)
	~ Hour + Hour <sup>2</sup>
	~ Julian date + Hour
	~ Julian date + Julian date <sup>2</sup> + Hour
	~ Julian date + Hour + Hour <sup>2</sup>
	~ Julian date + Julian date <sup>2</sup> + Hour + Hour <sup>2</sup>
Abundance	~ 3 Bioclimatics PCA axis + 3 Habitat cover PCA axis

HDS population size estimates were obtained by averaging retained secondary candidate sets models, based on their relative model performance using AICc (Fig 3). We excluded the Eurasian Sparrowhawk (*Accipiter nisus*), the Meadow Pipit (*Anthus pratensis*) and the Coal tit (*Periparus ater*), from model averaging and exclusively relied upon prediction from best final models owing to substantial differences observed among their secondary candidate sets models.

Model goodness-of-fit was assessed using an overdispersion coefficient metric ( $\hat{C}$ ; Johnson et al., 2010). We used the chi-square metric as the discrepancy measure between observed and expected counts. Computed  $\hat{C}$  corresponds to the ratio between the chi-square obtained from the fitted model to the mean of bootstrapped chi-squares obtained from simulated datasets based upon estimated parameters (Kéry & Royle, 2015). All models were fit according to a Poisson (P) distribution after top model assessment and calculation of  $\hat{C}$ , secondary candidate sets with  $\hat{C}$  top models exceeding 1.2 were calibrated using a Negative binomial (NB) distribution (Payne et al., 2018). For a global overview of our modelling approach, see Appendix S5. Out of 63 species, we excluded nine species from the analysis; three exhibited signs of underdispersion with  $\hat{C}$  values less than 0.9 while six had  $\hat{C}$  values exceeding 1.5 (Payne et al., 2018), showing signs of overdispersion, despite being calibrated using a negative binomial distribution, see table S2.1 for more details.

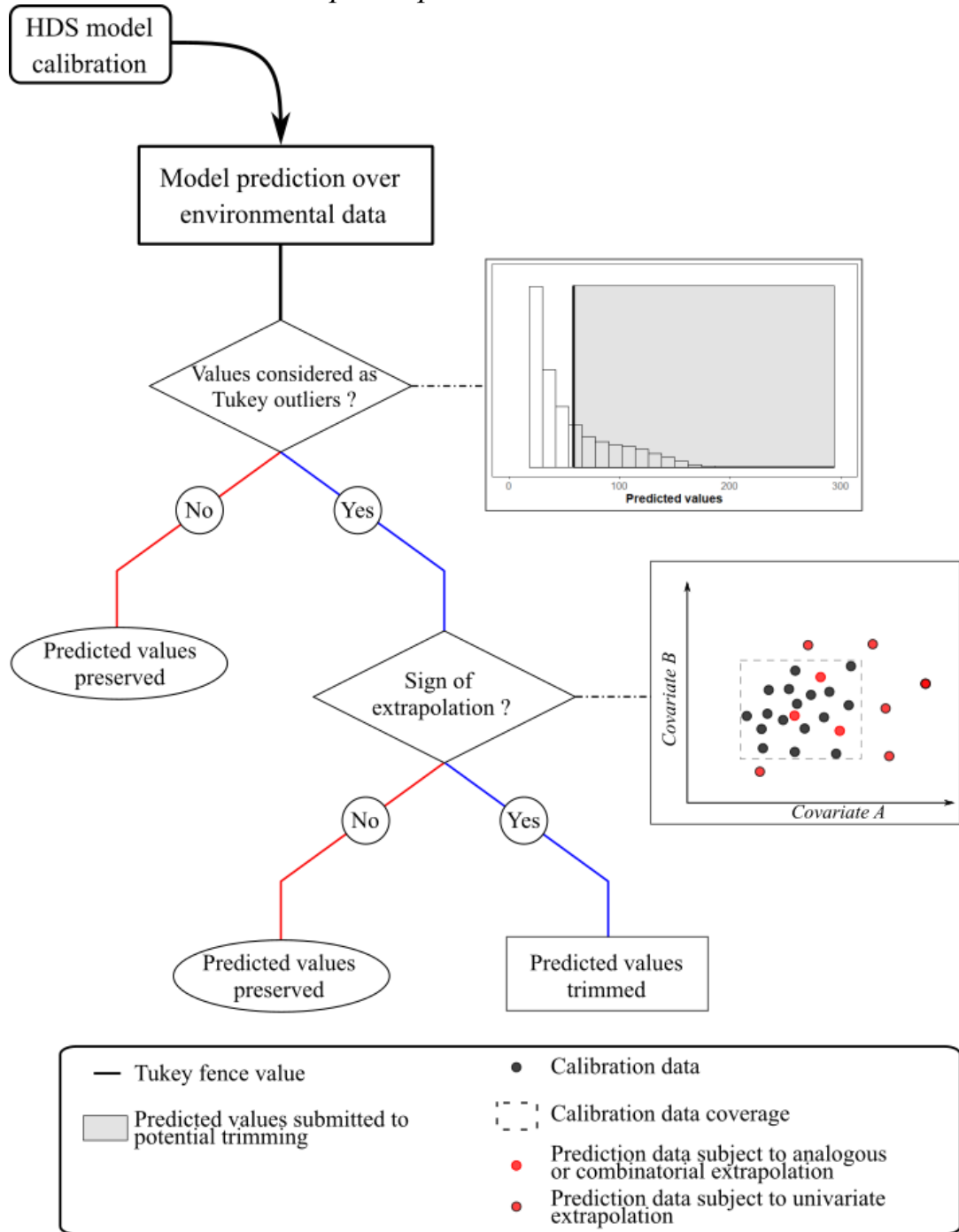
We assessed the robustness of our estimations to the exclusion of one year of data, corresponding to a third of the global dataset. We compared population size estimates from EPOC-ODF data collected over 2021-2023 to estimates obtained from EPOC-ODF data collected over 2021-2022. Using the 2021-2022 subset, we estimated the population sizes of 30 species, detected in at least 150 distinct sites. From these 30 species, seven mean population sizes estimated using 2021 to 2023 data were outside the confidence intervals estimated from 2021 to 2022 data, with a slightly smaller population size estimated (Table S2.2) overall highlighting robust estimations.

#### *Trimming of HDS population size estimate: assessment of model extrapolation*

Population sizes were obtained by summing predicted values over the prediction grid. As we intend to predict over a large surface, novel environmental conditions may arise,

leading to possible dissimilarities between the environmental gradient collected at survey sites and the environmental gradient over novel conditions (Elith & Leathwick, 2009). Mesgaran et al., 2014 described two types of extrapolation; (i) Novelty type I (NT1) where projected points (*i.e* prediction grid) are outside the range of individual covariates collected by the sampling scheme and (ii) Novelty type II (NT2) depicting the case when projected points are within univariate range but constitute novel combinations between covariates.

We trimmed predicted values (Fig 4) over prediction grid cells showing signs of NT1 extrapolation, to a threshold value determined using a Tukey fence (Tukey, 1977), estimated from the distribution of predicted values with  $k=1.5$ . Extrapolation assessments were done using *dsmextra* 1.1.5 R package (Bouchet et al., 2020). We used this post-prediction treatment to assess population size estimates stability. We measured the coefficient of variation, corresponding to the ratio between the standard deviation and the mean of the “untrimmed” and “outlier-trimmed” estimated range uncertainty. Large coefficients of variation imply great discrepancies in confidence intervals of untrimmed and outlier-trimmed estimated uncertainty intervals. This is mainly caused by the spatial filtering from the extrapolation assessment highlighting smaller geographic regions with similar environmental conditions from sampled ones and the trimming of predicted abundance outliers. Species with a coefficient of variation exceeding 30% were removed from the comparison of *ArGeom* and HDS population size estimates (Table S2.3).



**Figure 4 - Decision tree of the post-prediction treatment.** First, we analysed the distribution of predicted abundance values across the prediction grid and detected outliers, using a hinge of  $k = 1.5$  (Tukey, 1977). We compared each environmental condition of the prediction grid cell with the environmental condition collected by the sampling scheme and used for model calibration. When a prediction cell depicted signs of NT1 extrapolation and its estimated abundance was considered as an outlier, we trimmed the predicted grid cell abundance to the Tukey fences value.

For comparable estimates between *ArGeom* and HDS approaches, we restricted the prediction grid area species-wise for HDS estimation according to the distribution of their known breeding locations, collected over a 10x10km grid during the previous French atlas (Issa & Muller, 2015). To estimate breeding populations of species for which male identification was possible, either male vocalisations or visual distinctions because of sexual dimorphism, an ad-hoc filter was applied (Table S2.1), resulting in HDS estimates reflecting the male counts for those species.

As the *ArGeom* approach estimated species bird population sizes as a number of breeding pairs (Roché et al., 2013), for species where male identification in the field was impossible (no sexual dimorphism), we used all available data, after applying the phenological filter, and divided HDS estimates by two for comparable estimates with *ArGeom* population sizes. After retrieval of *ArGeom* estimates from the previous atlas (Issa & Muller, 2015), we updated these estimates using recent population trend estimates derived from the French breeding bird survey (FBBS; Jiguet et al., 2012) data spanning 2012-2023 (Table S2.1). Given the absence of a mean estimate in the *ArGeom* approach, we approximated it using the midpoint between the maximum and minimum estimated (Fig 5).

To study the differences between the two approaches, we measured  $\delta_{mean}$  corresponding to the percentage of the difference between HDS and *ArGeom* estimates.

$$\delta_{mean} = \frac{(Estimates_{ArGeom} - Estimates_{HDS})}{(Estimates_{ArGeom} + Estimates_{HDS})/2}$$

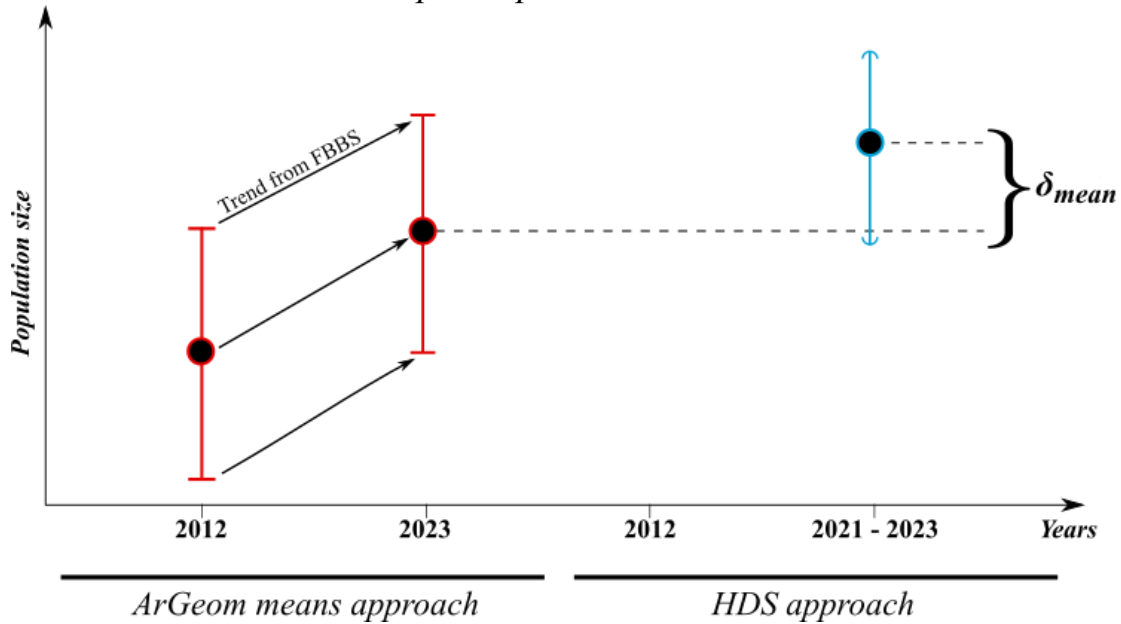


Figure 5 - Population size estimates comparison methodology, *ArGeom* in red and HDS in blue. *ArGeom* estimates for 2012 were obtained through calculations using both arithmetic and geometric means (Roché et al., 2013) and were then updated to 2023 using specific mean trend estimates from FBBS (Table S2.1). Flat intervals signify min and max value estimates, while curved intervals signify confidence intervals.

### *Study of variation of estimated population sizes between the two approaches*

As species detectability stems from physical traits and vocalisations, phylogenetic related species tend to have the same detectability (Johnston et al., 2014; Sólymos et al., 2018). We calibrated a Phylogenetic Generalised Linear Mixed Model (PGLMM) using the `phylglm` 1.0 (Li & Bolker, 2019) R package. We study  $\delta_{mean}$  variations across species while implementing a random effect covariance structured based on phylogenetic relatedness using phylogenetic distances retrieved from Burleigh et al., (2015). The PGLMM model was calibrated using (i) extracted detection probabilities from the availability state estimated through HDS (fig 3 and appendix S5) after model averaging of the final candidate sets models in regards to AICc scores, and (ii) *ArGeom* uncertainty as fixed variables. For *ArGeom* uncertainty, corresponding to the difference between maximal and minimal estimated values, we relied on the decimal logarithm to limit variation in  $\delta_{mean}$  solely due to different population size magnitudes.

Response weights consisted of normalised weights from the inverse of uncertainty around FBBS trends between 2012 and 2023 (Table S2.1), divided by the mean to limit excessive weight attribution and facilitate model convergence.

## Results

### *Species trends over 2012-2023*

From 2012 to 2023, out of 63 bird species, 15 showed a significant decrease ( $x = -22.79\% \pm 14.84$ ) in total population size, while 16 showed a significant increase ( $x = 28.02\% \pm 22.52$ ; see appendix S2 for species-related FBBS trends).

### *HDS population size estimations*

Out of the 54 species with acceptable values of overdispersion ( $\hat{C}$ ) using the *HDS* approach, we excluded eight species showing large discrepancies in population size estimates (Appendix S2, Table S2.3) between pre- and post-prediction treatment (Fig. 2.).

Out of the remaining 46 species used for comparison between *ArGeom* and *HDS* estimates, *HDS* models showed acceptable values of overdispersion ( $\hat{C}$ ) ranging from 0.94 to 1.2 ( $\bar{x} = 1.07 \pm 0.06$ ) for 38 species calibrated using a Poisson distribution and 1.09 to 1.47 ( $\bar{x} = 1.27 \pm 0.13$ ) for eight species calibrated using a Negative binomial distribution.

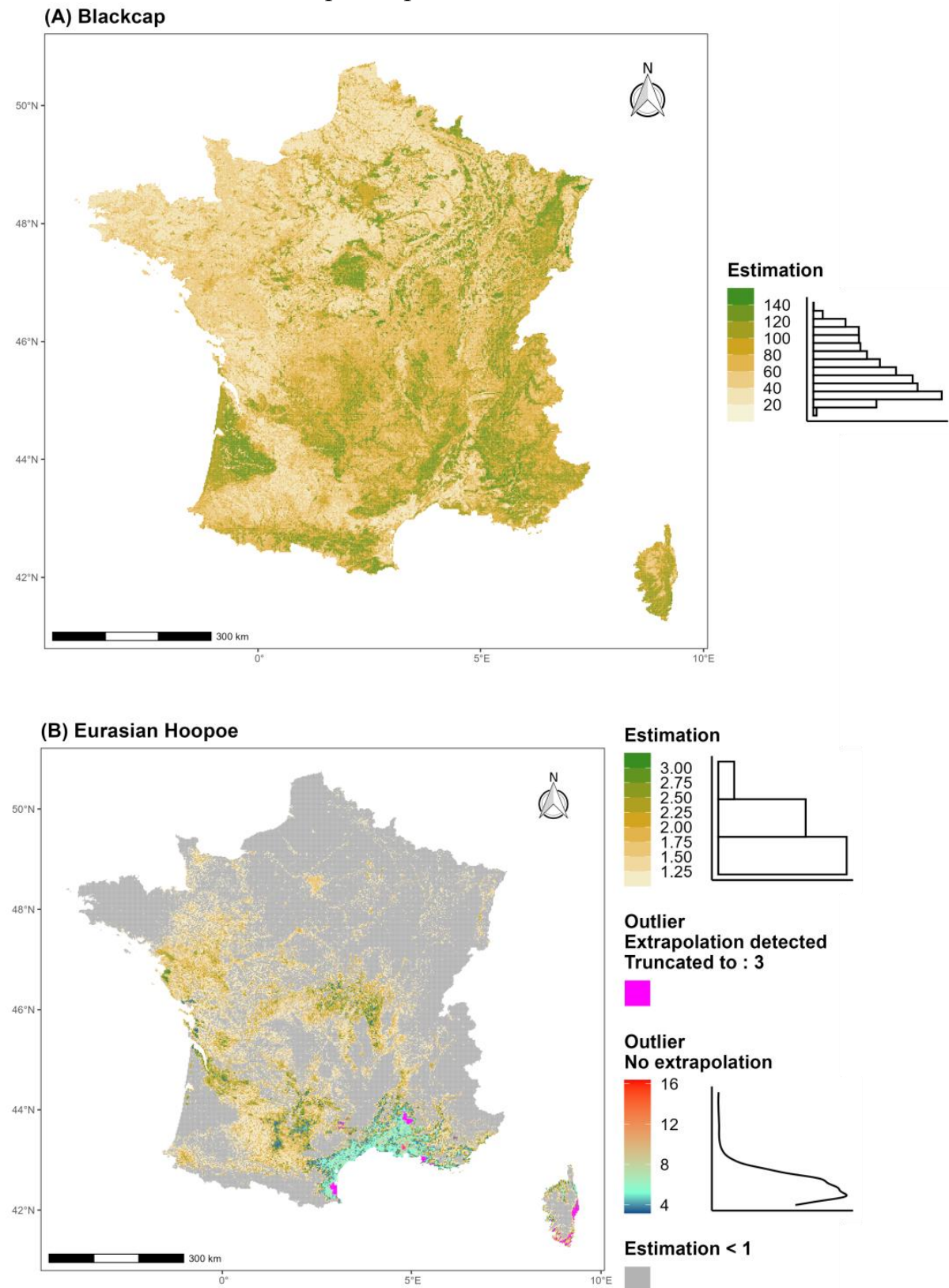


Figure 6 - Examples of resulting abundance maps for the (A) Blackcap (*Sylvia atricapilla*), and (B) Eurasian Hoopoe (*Upupa epops*). Estimations correspond to the number of male, or potential pairs (see Table S2.3)

over a 4 km-squared area. Grid cell predictions are categorised into three groups: (1) those with estimated abundance not considered as outliers depicted with a green colour step gradient and its distribution histogram ; (2) those with estimated abundance considered as outliers and not subject to NT1 extrapolation, displayed in a blue-to-red gradient, along with their distribution density; and (3) those with estimated abundance considered as outliers with novel environmental conditions subject to NT1 extrapolation highlighted in pink with the Tukey value used for trimming in the post-prediction treatment.

### *Population size comparison between ArGeom and HDS*

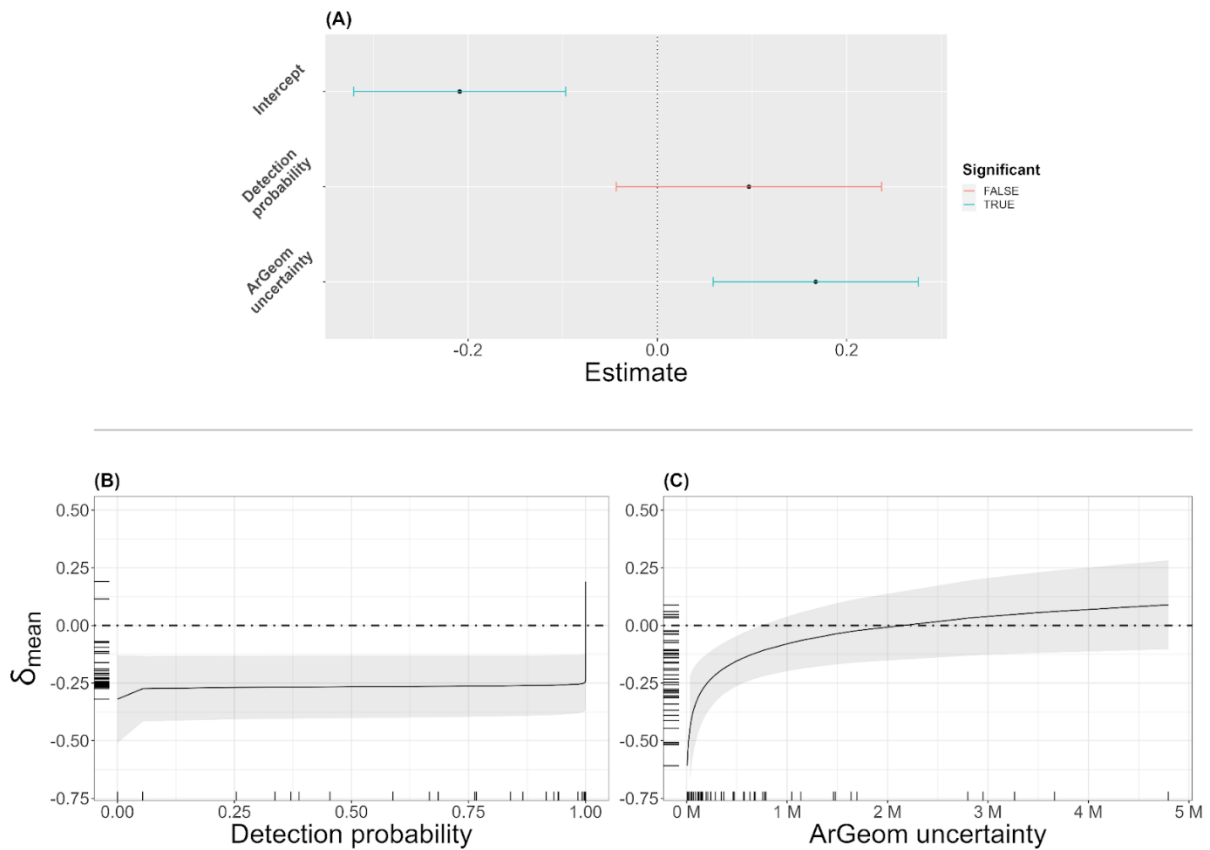
Across all species, estimated mean density ranges from 0.09 to 27.51 individuals per square kilometre, while ArGeom range uncertainty varies from 3.9 to 6.69 on the decimal logarithm scale corresponding to variations from 7,920 to 4,850,000 in estimated number of pairs. See Table S6.1-3 for more details about species estimated population size according to *ArGeom* and HDS approaches.

A comparison between updated *ArGeom* and HDS estimated population sizes showed that HDS estimates were higher than *ArGeom* for 30 of the 46 species tested (Table S6.1). Our results suggest lower estimates from *ArGeom* ( $\delta_{mean} < -0.2$ ), either for open habitat specialists such as European Stonechat (*Saxicola rubicola*), European Goldfinch (*Carduelis carduelis*) or Eurasian Linnet (*Carduelis cannabina*) than for forest generalists such as Great Spotted Woodpecker (*Dendrocopos major*), Blue Tit (*Cyanistes caeruleus*) or Eurasian Blackcap (Fig 6.A) with  $\delta_{mean} = -0.629 \pm 0.4$ , over 22 species. Species whose estimations were similar ( $\delta_{mean} \in [-0.2; 0.2]$ , with on average  $\delta_{mean} = -0.006 \pm 0.096$ , over 17 species) between *ArGeom* and HDS included species such as Eurasian Wren (*Troglodytes troglodytes*), Great Tit (*Parus major*), European Robin (*Erithacus rubecula*) and European Nuthatch (*Sitta europea*). Fewer species, mainly characterised by greater maximal observation distances, such as Common Cuckoo (*Cuculus canorus*), Eurasian Hoopoe (Fig 6.B) and European Blackbird (*Turdus merula*) had higher population sizes estimated by

over 7 species; see Appendix S6 for population size comparison table).

Results from the PGLMM (Fig 7.A) showed an overall significantly lower ArGeom population size estimates (-0.209, with 95% CI : [-0.321;-0.097], Pval = 0.001), as well as a significant positive effect of ArGeom range intervals (0.167, with 95% CI : [0.059;0.276], Pval = 0.003) on the differences between the two approaches. Species detection probabilities had no significant effect (0.097, with 95% CI : [-0.043;0.237], Pval = 0.176) on  $\delta_{mean}$  variation.

Marginal effect plots from the PGLMM model showed that the mean response of  $\delta_{mean}$  over species detection probability was predominantly negative, ranging from -0.45 to -0.1 (Fig 7.B), for ArGeom uncertainty. This showed that  $\delta_{mean}$  tended towards the convergence of population size estimates (Fig 7.C) for species with larger estimated interval ranges. There were no signs of multicollinearity (VIF < 5; James et al., 2013) between the two variables.



**Figure 7 - Results from the PGLMM. (A) Confidence intervals of the model coefficient, coefficients significantly different from 0 are represented in blue. Marginal effect plots of population size estimate differences ( $\delta_{mean}$ ) responses.  $\delta_{mean}$  responses are predicted over gradients of focal terms, either species detection probabilities (B) or *ArGeom* reported uncertainties (C), while other covariates are held constant at their mean. Species detection probabilities and *ArGeom* uncertainty are represented on their natural scales, after inverse logit and inverse decimal logarithm transformation, respectively. Dot-dash line corresponds to a  $\delta_{mean}$  of 0, signifying estimated population size convergence by the two approaches, negative and positive values of  $\delta_{mean}$  reflect lower and higher population size estimates of *ArGeom* relative to those obtained using HDS.**

## Discussion

Our results showed that bird population size estimates from the previous *ArGeom* approach, not accounting for the observation process nor habitat affinity covariates, are predominantly lower than population sizes estimated from the HDS approach, up to 65% of species. While we found that the prior estimated uncertainty ranges from *ArGeom* had a positive effect on the convergence of population sizes estimated by the two methodological approaches (expert estimates based on atlas data vs predicted estimates from modelled citizen science data accounting for detection probabilities), we did not find a significant effect of

species detection probabilities which could explain the differences between the two approaches. We show that ArGeom produces population sizes that are largely lower than those obtained by HDS ( $\delta_{mean} < 0.2$ ) regardless of habitat specialisation or affinity. This is likely due to the methodology used for ArGeom that did not account for the detection process nor for species-habitat relationships when extrapolating locally known abundances to unsampled locations.

Despite the lack of significant evidence for the effect of species detection probabilities on population size estimations, our results tend to corroborate previous studies where ignored detection processes had likely biased ecological inferences, including species distribution models (Kéry et al., 2010), population trends (Norvell et al., 2003; Schmidt et al., 2013) and population sizes (Kéry et al., 2005). This lack may be especially true in a context of global change, where avian breeding phenology showed evidence of shifts towards earlier breeding over the years (Parmesan, 2007; Devictor et al., 2012; Gaüzère & Devictor, 2021) to synchronise with their food sources (Visser et al., 2006; Michel et al., 2016). Such shifts induce diverse species-related seasonal and inter-annual changes in detectability that need to be accounted for, particularly for schemes spanning over multiple species (Lehikoinen, 2013).

In France, the ArGeom approach was developed in the context of moderate semi-quantitative data collection with an acknowledged uneven participation across the territory (Roché et al., 2013; Issa & Muller, 2015). The semi-quantitative data collection was based on the estimation of the number of breeding pairs over 10x10km grids derived from a mixture of count prospects and expert opinions collected over 1953 out of 5879 10x10km grids (Issa & Muller 2015). As the primary goal of this approach was to give a likely magnitude of population size across the territory (Roché et al., 2013), ArGeom intervals were produced by

the extrapolation from the initially prospected 10x10km grids to all metropolitan grid cells considered suitable for breeding, after the detection of potential and confirmed breeding evidence. Although proved useful and relevant to assess population sizes when large-scale quantitative data on species occurrences are lacking or are unevenly distributed, such a methodology implies greater uncertainty intervals for abundant and broadly distributed species and smaller intervals for scarce and narrowly distributed species when not accounting for measurement uncertainty.

Contrary to generalist species, which have widespread distributions due to broader habitat niche breadths, specialists are generally more localised (Clavel et al., 2011) and typically use a smaller range of habitats (Julliard et al., 2006). Despite a growing interest in rare species-focused monitoring (Fontaine et al., 2022), citizen sciences programs are mainly designed for large-scale multi-species surveys (Devictor et al., 2010). Citizen science schemes balance between a trade-off among data quantity and data quality, corresponding to either the acquisition of a great quantity of unstructured scheme, or the acquisition of standardised data implying replicated visits over randomly sampled locations (Devictor et al., 2010). As such, in the first case, citizen science schemes could be more prone to false-negative errors, resulting in biased inferences over habitat cover relations due to omission of the detection process (Johnston et al., 2022). In the second case, given the small habitat range of specialist species and the scale of the territory sampled (e.g., here the metropolitan French territory), citizen science sampling schemes could be representative of the entire territory sampled but with a higher risk of missing some key habitats and associated specialist species.

A recent study about long-term effects of climate and land use changes on bird communities (Gaüzère et al., 2020) showed that both generalist cold-dwelling species, such as the Common Chiffchaff (*Phylloscopus collybita*) or the Eurasian Blackcap, and warm-dwelling species, such as the Common Nightingale (*Luscinia megarhynchos*) had the most substantial negative and positive contributions to the trend in Community Thermal Index (CTI), a community-weighted index representing the realised thermal niche of a community based upon species relative abundance and species thermal indices (STI). In the present study, these species tended to have lower population sizes estimated when the detection process was omitted compared to estimates based on our modelling approach. As a result, this could affect the estimations of their contribution to the calculation of community-weighted mean indices, such as CTI, and therefore bias the estimation of the trend in community thermal response and subsequent studies of aggregated indices, which are known to display large regional variation (Rigal & Knape, 2024). We, therefore, suggest that considering the detection process in studies relying on community-weighted indexes by species' relative abundances could be as important as it is for estimating population sizes.

Community indices such as species diversity (Ricotta, 2005) and functional diversity (Villéger et al., 2008; Gaüzère et al., 2019) are commonly use species relative abundance as a basis, without taking into account the detection process (Pillar & Duarte, 2010), despite multiple studies showing it could affect community indices inference (Tingley & Beissinger, 2013; McNew & Handel, 2015; Jarzyna & Jetz, 2016; Richter et al., 2021).

Our study also suggested that lower or higher population sizes estimated from ArGeom were not randomly distributed among species according to their conservation status. Out of the 46 species estimations used in the comparison analysis, 10 had an unfavourable conservation status in France (*i.e.* lower than Least Concern, LC; UICN France et al., 2016).

Among these species of conservation concern, two species, European Greenfinch (*Carduelis chloris*) and European Turtle Dove (*Streptopelia turtur*), showed no signs of difference in their population sizes. By contrast, five species, European Stonechat, Barn Swallow (*Hirundo rustica*), Red-backed Shrike (*Lanius collurio*), Eurasian Kestrel (*Falco tinnunculus*) and Willow Warbler (*Phylloscopus trochillus*) considered as NT (Near Threatened) and three species, Eurasian Linnet, European Goldfinch and European Serin (*Serinus serinus*) considered as VU (Vulnerable) had lower population sizes estimated from ArGeom than from HDS approach (NT :  $\bar{\delta}_{mean} = -0.608 \pm 0.217$  and VU :  $\bar{\delta}_{mean} = -0.667 \pm 0.146$ ). Our results showed that these species may need a reevaluation of their conservation status and highlight the need to rely on hierarchical models taking account of the detection process in ecological inferences, given that potential misclassification of population conservation status may arise from process noise and observation error (Connors et al., 2014). As conservation policy decisions depend on uncertainty levels (Williams, 2003; Freckleton, 2020), assessing measurement error through the integration of the detection process (Nichols et al., 2011) could provide more reliable ecological inferences (Guillera-Arroita et al., 2014). CS schemes are becoming more and more a reliable source of data to ensure biodiversity monitoring (Chandler et al., 2017) and can, through standardisation (Buckland & Johnston, 2017; Johnston et al., 2019), contribute to the calibration of data-hungry models such as

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives* Article  
hierarchical models for reliable ecological inferences (Isaac et al., 2020; Kéry & Royle, 2021; Johnston et al., 2022).

### *Comparison to other European countries*

Another way to assess the relevance of the two estimation approaches would be to compare their population size estimates to the ones obtained from other European countries, using a ratio between countries to produce comparable estimates. Such an approach should however be used with caution because it would be limited by comparability in habitat repartitions or biogeographical considerations among different European countries. To go further into inter-country comparisons, we relied on the German population sizes estimated for the previous European Bird Directive (BirdLife International, 2021) obtained from both point count and territory mapping methods (Gedeon et al., 2015). For abundant species such as the Blackcap ( $\delta_{mean} = -0.29$ ; German population size expressed in millions of pairs = [7.17 - 9.49]), both approaches led to similar results than German population estimates, while HDS estimates were closer to German population sizes for the Firecrest (*Regulus ignicapilla*;  $\delta_{mean} = -0.46$ ; [1.92 - 2.85]) and the Blue Tit ( $\delta_{mean} = -0.42$ ; [5.01 - 7.41]). For species with higher population sizes estimated by ArGeom than HDS ( $\delta_{mean} > 0.2$ ), the Common Cuckoo ( $\delta_{mean} = 0.25$ ; [0.58 - 0.95]) and the Corn Bunting (*Emberiza calandra*;  $\delta_{mean} = 0.6$ ; [0.25 - 0.44]) showed estimates of German populations closer to the HDS than the ArGeom approach. Finally, for the Common Whitethroat (*Curruca communis*;  $\delta_{mean} = 0.42$ ; [0.93 - 1.47]), the German population size is closer to ArGeom estimates (see appendix S6.4 for additional information).

Regarding magnitudes, both approaches produced similar estimates compared to German ones. However, due to different sampling and modelling methods used, these formal comparisons, although informative, need to be more fully satisfying and highlight the discrepancies in sampling and analytical methods across the European continent (Keller et al., 2020). Such differences could be accounted for, either by (i) a global standardisation of schemes, as promoted by the PECBMS (Pan-European Common Bird Monitoring Scheme; Brlík et al., 2021) for species trends, but also (ii) through the use of Integrated Models (IM) capable of mobilising data from multiple and somewhat heterogeneous sources (Isaac et al., 2020; Zipkin et al., 2021a).

### *Study limitations*

Our approach relies on data collected from the EPOC-ODF structured CS schemes, providing data with repeated visits. However, as is, the frequentist framework of unmarked R package (Kellner et al., 2023) does not permit inferences on social species occurring in large flocks. Taking account of social species during the breeding season (corresponding to 1/10th of the scheme targeted species) would therefore require a Bayesian framework to include the effect of flock size on species detectability (Clement et al., 2017).

Given the timeframe and the sampling design, *i.e.*, all sites are not visited every year to maximise the number of total surveyed locations, it is not possible to estimate species demographic parameters, such as survival and recruitment (Sollmann et al., 2015; Schmidt & Rattenbury, 2018). We also assumed a sex ratio of 1:1 for species without sexual dimorphism, during the breeding season, which could potentially bias estimates for species deviating from this assumption. Taking account of species population structure requires

frameworks such as Integrated Population Models (IPM; Schaub & Ullrich, 2021) and specific data collection (King, 2014), for instance, bagging or nest surveillance.

As obtaining relevant predictions of species abundance over unsampled environmental conditions was one of our main methodological challenges, we used environmental data condensing habitat information (Tredennick et al., 2021). To fit our statistical framework, we assumed that most bird species would interact with their habitat following a linear relationship (see Fig. 3). We therefore used PCA reduction to summarise species linear responses to national-scale habitat gradients including forest-to-open-field cultures, open-field cultures-to-pastures and perennial crop-to-urban habitats (see appendix S3). PCA reduction permits model convergence by condensing complex habitat structures to a small number of environmental covariates, though it could bias estimates of species thriving in a specific habitat restricted to the extreme edge of the sampled gradients. Other methods such as Spatially Varying Covariates models (SVC; Gelfand et al., 2003) could be used to better account for habitat structure complexity across spatial gradients (Thorson et al., 2023).

Previous studies have shown that unaccounted variations in species availability, considering a constant detection probability or unmodelled variations, could lead to substantial bias in estimated abundance (Link et al., 2018; Barker et al., 2018; Duarte et al., 2018). N-mixture biased estimations can be linked to non-assessment of the sampled area, where a smaller or greater sampled effective area could lead to under- or overestimation (Kéry & Royle, 2015). In our study, as we relied on distance sampling methods, we define an effective sampled area, based upon collected observation distance, but we also assumed that individuals considered exposed to the sampling (*i.e.* 'statistically' available for modelling) could still be undetected due to small species home ranges or plot-specific habitat cover (Chandler et al., 2011; see Table 1 and Figure 2, for covariates used to model species

detectability and appendix S5 for model formulation). Despite such consideration, for the HDS model, we assumed that detected individuals were homogeneously distributed over the sampled area. Violating this assumption could lead to within-sample variation that needs to be accounted for, otherwise leading to biased estimates (Mizel et al., 2018).

Another potential drawback relies on the quantity of data collected through this structured CS scheme. Over the same breeding season, the semistructured scheme EPOC without temporal replicates nor fixed location requirements collected three times the amount of complete checklists as the structured EPOC-ODF scheme, highlighting CS trade-off of scheme standardisation upon data collection over spatial and temporal scales (Devictor et al., 2010). One way to address this trade-off would be to apply data integration methods mobilising multiple data sources to be used for ecological inferences (Zipkin et al., 2019, 2021b), either by estimating abundance of less recorded species through trait-based associations (Callaghan et al., 2021, 2022; Robinson et al., 2022) or by constructing joint likelihood functions (Fithian et al., 2015; Fletcher et al., 2019).

## Conclusions

Our results suggested an overall lower population size estimate of French common breeding birds obtained from the last French Breeding Bird atlas methodology than from the Hierarchical Distance Sampling modelling used in the present work. Using large-scale datasets from citizen science obtained from standard scheme initiatives allowed us to infer the variation in species abundance, while explicitly modelling the detection process separately from the ecological one. Not accounting for the observation process might have resulted in misleading expert-only estimations of population sizes in the previous atlas, at

least for some widespread species not benefitting from exhaustive surveys. In conclusion, our results advocate for more reliance on the use of statistical tools accounting for the detection process, such as hierarchical models, which, in association with large-scale citizen science data, could constitute a standard methodology to estimate reliable abundance from breeding bird atlases or biodiversity surveys deployed at national or geographically broader scales

### Acknowledgements

EPOC-ODF scheme is supervised by LPO (Ligue pour la Protection des Oiseaux), the representative of Birdlife in France. We thank all participants and local coordinators from regional instances contributing to data collection through [faune-france.org](http://faune-france.org), a collaborative online database.

HDS models were calibrated on the SACADO MeSU platform at Sorbonne Université at Paris – France, while environmental data were extracted using 2AD PCIA platform at Muséum Nationale d'Histoire Naturelle, Paris – France. We would like to thank Emmanuelle Porcher, Laurent Godet, Karine Princé and Jean-Yves Barnagaud for their guidance over this work, Lise Bartholus for the computation of the 2012-2023 FBBS trends and Cyril Eraud for his comments over model calibration; as well as two anonymous reviewers for their suggestions that greatly improved this work. A preprint version of this article has been peer-reviewed and recommended by PCIEcology (<https://doi.org/10.24072/pci.ecology.100683>).

### III.C Supplementary information

#### *S1: Scheme standardization*

EPOC-ODF sites are selected from a two-stage random sampling. In order to standardise survey effort over France and help local coordinators site assignment, ten 10x10km atlas grids are randomly selected by districts. Each selected grid is subsequently divided as twenty-five 2x2km grids. Ten out of twenty-five 2x2km grids are selected and categorised as "primary" and "secondary" sites (Fig S1.1). Participant birders are tasked to survey over the centroids of five of these selected sites.

Participants are asked to survey designated "primary" sites, but in regards to terrain specificities, inaccessibility or private land, they can survey one of the "secondary" selected sites. Each completed 10x10km grid cell is removed from the sampling pool for future years of the atlas survey.

Observers are asked to survey the selected sites three times during the breeding season, from March to June, with at least 30 days between each visit. Time periods for the first visits - 1st to 31th of march ; second visits - 1st of april to 8th of may and third visit - 9th of march to 15th of June (<https://oiseauxdefrance.org/get-involved/epoc-odf>). In high altitude, the first visit, typically in March, isn't compulsory and the consequent visits are adjustable in regards to snow conditions and site accessibility.

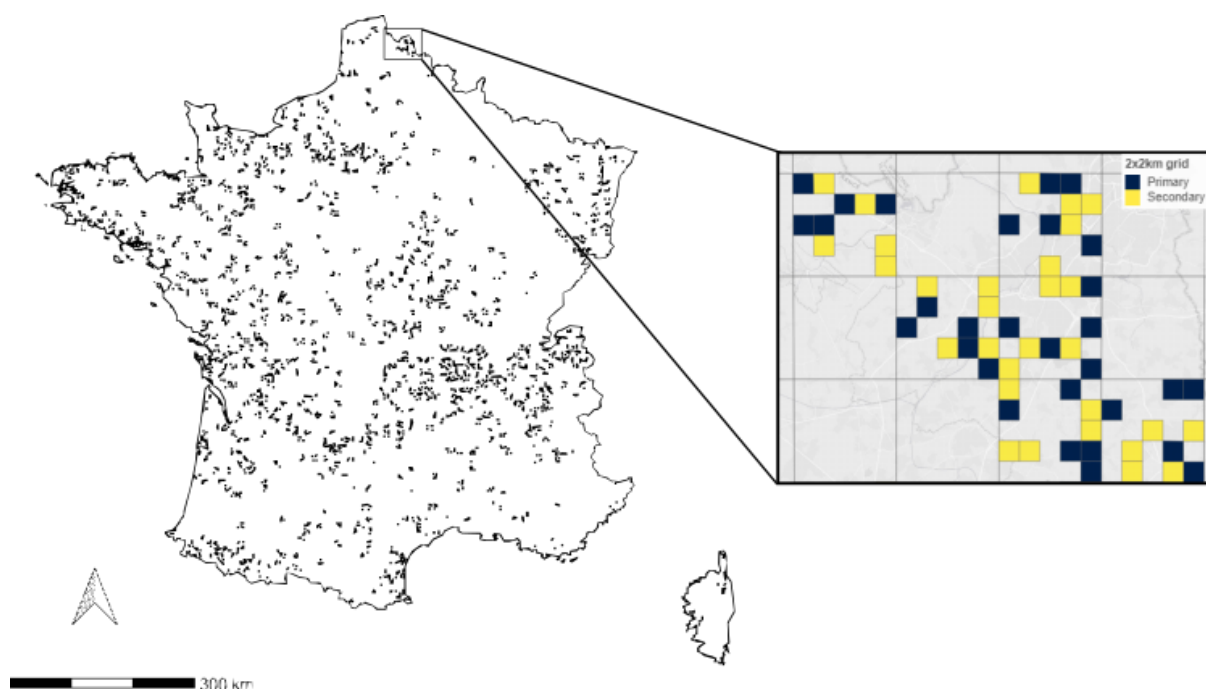


Figure S1.1 - Schematic of the systematic grid for EPOC-ODF sites.

S2: Species supplementary tables

*Species metadata*

Table S2.1 - Species metadata of 63 considered species regrouping FBBS trends between 2012-2023 with their confidence intervals, insignificant trends as well as significant increase and decrease are represented with beige, green and orange colours ; starting and ending phenological periods in julian date ; a priori data filtering based on identifiability of males during survey ad-hoc filters choices;  $\hat{C}$  assessment acceptance.

Species	FBBS trend (%) 2012-2023	Phenological filtering dates		A priori Male filtering	$\hat{C}$
		Start	End		
Barn Swallow	-13.4 [-20 ; -6.2]	122	182	FALSE	TRUE
Black Redstart	-10 [-15 ; -4.7]	92	152	TRUE	TRUE
Black Woodpecker	18.8 [4.8 ; 34.7]	61	152	FALSE	TRUE
Cetti's Warbler	132.2 [85.3 ; 190.8]	75	152	FALSE	FALSE
Cirl Bunting	14.7 [4.2 ; 26.3]	75	152	TRUE	TRUE
Coal Tit	-48.5 [-55.8 ; -40]	92	152	FALSE	FALSE
Common Buzzard	-4.8 [-11.1 ; 2]	61	152	FALSE	TRUE
Common Chaffinch	-8.3 [-10.9 ; -5.6]	92	152	TRUE	TRUE
Common Chiffchaff	-1.2 [-4.7 ; 2.3]	92	152	FALSE	TRUE
Common Cuckoo	-7.4 [-12.7 ; -1.7]	122	182	FALSE	TRUE
Common Nightingale	-4.7 [-9.8 ; 0.7]	122	182	FALSE	TRUE
Common Raven	65.6 [20.7 ; 127.1]	61	152	FALSE	FALSE
Common Redstart	14.6 [1.8 ; 29]	122	182	TRUE	TRUE
Common Wood-Pigeon	46.3 [40 ; 52.9]	92	152	FALSE	TRUE
Corn Bunting	21 [8.2 ; 35.4]	75	152	FALSE	TRUE
Crested Tit	-28.5 [-38 ; -17.5]	92	152	FALSE	TRUE
Dunnock	-24.1 [-29.9 ; -17.8]	92	152	FALSE	TRUE
Eurasian Blackbird	-6.8 [-10 ; -3.5]	75	152	TRUE	TRUE
Eurasian Blackcap	8.8 [5.5 ; 12.1]	106	152	TRUE	TRUE
Eurasian Blue Tit	9.1 [3.4 ; 15]	92	152	FALSE	TRUE
Eurasian Bullfinch	-48 [-56.4 ; -38.1]	75	152	FALSE	TRUE
Eurasian Collared-Dove	-4.6 [-9.8 ; 0.8]	75	152	FALSE	FALSE
Eurasian Golden Oriole	15.6 [6.9 ; 25]	122	182	TRUE	TRUE
Eurasian Green Woodpecker	-7.8 [-13.3 ; -1.9]	61	121	FALSE	TRUE

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives* *Supplementary information*

Eurasian Hoopoe	-8.3 [-20.5 ; 5.8]		106	167	FALSE	TRUE
Eurasian Jay	-14.9 [-19.6 ; -9.9]		75	167	FALSE	FALSE
Eurasian Kestrel	7.3 [-1.8 ; 17.2]		92	152	FALSE	TRUE
Eurasian Linnet	34.7 [18.7 ; 52.8]		106	167	TRUE	TRUE
Eurasian Magpie	11.5 [3.5 ; 20]		61	121	FALSE	FALSE
Eurasian Nuthatch	-31.8 [-37.6 ; -25.6]		75	152	FALSE	TRUE
Eurasian Skylark	-9.6 [-14.2 ; -4.7]		92	152	FALSE	TRUE
Eurasian Sparrowhawk	-12.6 [-28.5 ; 6.8]		75	152	FALSE	TRUE
Eurasian Wren	-15 [-18.6 ; -11.1]		92	152	FALSE	TRUE
European Goldfinch	13.7 [4.7 ; 23.6]		106	167	FALSE	TRUE
European Greenfinch	-42.5 [-46.8 ; -37.9]		106	167	TRUE	TRUE
European Robin	59.8 [52.9 ; 67]		75	152	FALSE	TRUE
European Serin	-3.7 [-14.1 ; 7.9]		92	152	FALSE	TRUE
European Stonechat	23.1 [12.3 ; 34.8]		92	152	TRUE	TRUE
European Turtle-Dove	-25.9 [-32.1 ; -19.1]		136	182	FALSE	TRUE
Firecrest	15.6 [4.6 ; 27.8]		92	152	FALSE	TRUE
Goldcrest	-46.1 [-53.4 ; -37.7]		92	152	FALSE	TRUE
Great Spotted Woodpecker	15 [8.8 ; 21.4]		61	152	FALSE	TRUE
Great Tit	-1.6 [-5.1 ; 2]		92	152	FALSE	TRUE
Greater Whitethroat	-5.6 [-11.6 ; 0.9]		122	182	FALSE	TRUE
Hawfinch	-17.5 [-33.5 ; 2.4]		75	152	FALSE	TRUE
Marsh Tit	-47.7 [-54 ; -40.4]		75	152	FALSE	TRUE
Meadow Pipit	-41.4 [-62.8 ; -7.7]		106	167	FALSE	FALSE
Melodious Warbler	-18 [-24.4 ; -11]		122	182	FALSE	TRUE
Middle Spotted Woodpecker	53.1 [21.5 ; 93]		61	152	FALSE	FALSE
Mistle Thrush	-3.4 [-11.8 ; 5.8]		75	152	FALSE	TRUE
Red-backed Shrike	49.4 [32.3 ; 68.7]		136	182	FALSE	TRUE
Ring-necked Pheasant	-1.8 [-12.1 ; 9.7]		61	152	TRUE	TRUE
Short-toed Treecreeper	13.2 [6.3 ; 20.6]		75	167	FALSE	TRUE
Song Thrush	4.7 [-1 ; 10.7]		75	152	FALSE	TRUE
Stock Dove	102.7 [63 ; 152]		75	152	FALSE	TRUE
Tree Pipit	-38.9 [-45 ; -32.2]		122	182	FALSE	TRUE
Western Bonelli's Warbler	7.1 [-10.9 ; 28.6]		122	182	FALSE	TRUE
Western Yellow Wagtail	-13.7 [-27.5 ; 2.6]		122	182	FALSE	TRUE
White Wagtail	-4.9 [-12.9 ; 3.8]		92	152	FALSE	TRUE
Willow Warbler	-44.2 [-54.5 ; -31.6]		122	182	FALSE	TRUE
Wood Lark	89.6 [69.4 ; 112.3]		92	152	FALSE	TRUE
Yellowhammer	-49.4 [-53.5 ; -44.9]		92	152	TRUE	TRUE
Zitting Cisticola	515 [340.3 ; 758.9]		92	167	FALSE	FALSE

*Chapitre 1 : Estimation des tailles de population d’oiseaux communs nicheurs à l’aide de données de sciences participatives* *Supplementary information*  
**Robustness analysis**

The robustness analysis consisted of comparing population size estimated using a subset of the calibration dataset. For the period 2021-2022, EPOC-ODF scheme collected 21 662 completed checklists (~79% of total checklists during 2021-2023) over 2874 sites (~74% of total sites surveyed).

Out of the 30 species whose models showed no sign of under/overdispersion ( $\text{large}\hat{C}$ ), using 2021-2022 EPOC-ODF dataset, nine species had 2021-2023 population size estimates outside of the confidence intervals determined from 2021-2022. 28 species had their confidence intervals shrank when calibrated from the 2021-2023 dataset, while two species, the Eurasian Linnet (*Carduelis cannabina*) and the Cirl Bunting (*Emberiza cirlus*) had wider confidence intervals using the 2021-2023 dataset. The shrinkage of the confidence interval was estimated by comparing uncertainty ranges (up - low) of the two periods.

**Table S2.2 - Robustness analysis table comparing species population sizes estimated using data collected from 2021-2023 and 2021-2022 EPOC-ODF scheme. Species which 2021-2023 population size estimated were outside of the confidence interval estimated using 2021-2022 dataset are highlighted in bold. Underlined-italic species names indicate species showing signs of under/overdispersion (Table S2.1). Displayed population sizes correspond to estimations derived from the framework outlined in the present paper after outlier-trimming. Species without estimation using data collected by the EPOC-ODF scheme 2021-2022 correspond to species detected in less than 150 distinct sites.**

Species	Estimation ODF 2021-2023			Confidence interval shrinkage	Estimation ODF 2021-2022		
	low	mean	up		low	mean	up
<b>Eurasian Blackcap</b>	8.6M	9.5M	10.5M	TRUE	12.7M	14.0M	15.6M
European Robin	7.7M	8.5M	9.4M	TRUE	7.4M	8.3M	9.3M
<b>Common Chaffinch</b>	7.2M	7.9M	8.8M	TRUE	9.2M	10.2M	11.4M
Eurasian Blue Tit	4.7M	5.4M	6.3M	TRUE	4.5M	5.3M	6.3M
Great Tit	4.9M	5.4M	6.1M	TRUE	4.8M	5.4M	6.2M
<b>Eurasian Blackbird</b>	4.0M	4.5M	5.1M	TRUE	5.2M	5.8M	6.5M
European Goldfinch	3.0M	4.1M	5.5M	/	/	/	/
Common Wood-Pigeon	3.2M	3.6M	3.9M	TRUE	3.4M	3.8M	4.2M

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives*

*Supplementary information*

Common Chiffchaff	3.1M	3.5M	4.0M	TRUE	3.2M	3.7M	4.3M
Eurasian Wren	3.0M	3.4M	3.9M	TRUE	3.1M	3.6M	4.3M
Barn Swallow	1.4M	2.0M	2.9M	/	/	/	/
Eurasian Linnet	551.3K	1.7M	5.6M	FALSE	1.7M	3.2M	6.0M
<u>Coal Tit</u>	3.0M	1.7M	5.2M	/	/	/	/
Eurasian Skylark	1.5M	1.5M	2.1M	/	/	/	/
<b>Short-toed Treecreeper</b>	1.1M	1.3M	1.6M	TRUE	1.6M	2.0M	2.5M
Firecrest	962.6K	1.3M	1.8M	TRUE	855.9K	1.2M	1.7M
<u>Eurasian Jay</u>	1.1M	1.3M	1.5M	/	/	/	/
<u>Eurasian Magpie</u>	900.9K	1.2M	1.8M	/	/	/	/
Dunnock	895.6K	1.1M	1.5M	TRUE	971.1K	1.3M	1.8M
<b>Song Thrush</b>	966.4K	1.1M	1.2M	TRUE	1.4M	1.7M	2.2M
White Wagtail	713.4K	1.0M	1.5M	TRUE	895.6K	1.6M	2.8M
<b>Eurasian Nuthatch</b>	831.3K	1.0M	1.3M	TRUE	1.3M	1.9M	2.6M
Goldcrest	1.3M	980.6K	2.7M	/	/	/	/
Cirl Bunting	778.7K	968.5K	1.4M	FALSE	595.3K	798.0K	1.1M
Black Redstart	717.1K	966.4K	1.3M	TRUE	834.2K	1.1M	1.5M
Great Spotted Woodpecker	752.6K	930.2K	1.2M	TRUE	830.0K	1.0M	1.3M
<u>Eurasian Collared-Dove</u>	716.6K	903.1K	1.2M	/	/	/	/
European Greenfinch	666.0K	881.2K	1.2M	TRUE	604.7K	823.3K	1.1M
Common Nightingale	739.8K	874.9K	1.0M	/	/	/	/
<u>Meadow Pipit</u>	237.2K	777.3K	2.9M	/	/	/	/
Hawfinch	422.3K	747.4K	1.4M	/	/	/	/
European Stonechat	494.2K	730.1K	1.1M	TRUE	490.3K	783.9K	1.3M
European Serin	532.4K	714.8K	979.5K	/	/	/	/
Marsh Tit	440.1K	672.8K	1.1M	/	/	/	/
Greater Whitethroat	475.9K	613.5K	795.1K	/	/	/	/
Eurasian Sparrowhawk	30.2K	603.5K	7.3M	/	/	/	/
Melodious Warbler	379.2K	499.1K	662.2K	/	/	/	/
Crested Tit	312.2K	492.5K	840.5K	/	/	/	/
Yellowhammer	370.4K	453.6K	698.9K	TRUE	414.8K	518.7K	815.6K
Eurasian Bullfinch	269.1K	437.2K	1.1M	/	/	/	/
Red-backed Shrike	244.1K	419.9K	752.0K	/	/	/	/
<b>Common Buzzard</b>	270.6K	367.8K	502.2K	TRUE	380.5K	494.0K	641.7K
European Turtle-Dove	265.3K	344.1K	449.8K	/	/	/	/
Mistle Thrush	296.2K	341.4K	417.2K	TRUE	332.4K	406.5K	547.1K
Western Bonelli's Warbler	148.8K	322.5K	757.7K	/	/	/	/
Eurasian Golden Oriole	244.9K	317.7K	409.3K	/	/	/	/
<b>Eurasian Green Woodpecker</b>	263.2K	316.1K	380.8K	TRUE	337.4K	414.6K	509.1K

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives*

*Supplementary information*

Ring-necked Pheasant	196.1K	293.4K	434.3K	TRUE	176.6K	274.4K	420.1K
Western Yellow Wagtail	276.0K	249.9K	410.5K	/	/	/	/
Tree Pipit	105.4K	195.3K	355.6K	/	/	/	/
Corn Bunting	150.1K	194.4K	254.3K	/	/	/	/
Wood Lark	153.1K	192.9K	245.2K	TRUE	135.2K	201.1K	305.7K
Eurasian Kestrel	126.1K	192.5K	295.9K	TRUE	148.2K	233.0K	368.5K
<b>Common Cuckoo</b>	129.8K	162.3K	199.9K	TRUE	770.2K	1.7M	3.7M
Common Redstart	39.3K	129.8K	283.0K	/	/	/	/
Willow Warbler	34.5K	113.1K	341.6K	/	/	/	/
<u>Common Raven</u>	26.0K	104.4K	340.2K	/	/	/	/
<u>Cetti's Warbler</u>	54.3K	80.6K	124.2K	/	/	/	/
<u>Zitting Cisticola</u>	40.1K	64.1K	106.6K	/	/	/	/
Black Woodpecker	26.6K	44.0K	71.9K	/	/	/	/
Stock Dove	20.3K	42.0K	86.6K	/	/	/	/
Eurasian Hoopoe	25.3K	41.1K	64.5K	TRUE	31.7K	52.1K	83.0K
<u>Middle Spotted Woodpecker</u>	15.9K	31.1K	65.7K	/	/	/	/

*Chapitre 1 : Estimation des tailles de population d’oiseaux communs nicheurs à l’aide de données de sciences participatives* *Supplementary information*  
**Estimates stability**

We assessed the stability of the estimated population size, over previously known breeding locations (Issa & Muller, 2015), by measuring the coefficient of variation of the two uncertainty ranges, before and after application of the post-prediction process described in Fig 2. Out of the 54 species with convenable  $\hat{C}$  values, eight species: Eurasian skylark (*Alauda arvensis*), Cirl bunting (*Emberiza cirlus*), Western Yellow wagtail (*Motacilla flava*), Stock dove (*Columba oenas*), Goldcrest (*Regulus regulus*), Crested tit (*Lophophanes cristatus*), Eurasian Bullfinch (*Pyrrhula pyrrhula*) and Yellowhammer (*Emberiza citrinella*) showed large discrepancies ( $CV > 30\%$ ) between the two estimates.

**Table S2.3 - A comparison of the 54 species, with converging HDS model (Table S2.1), population size estimates before and after trimming of extreme predicted abundance values, as shown in Figure 2, and the coefficient of variation (CV) calculated using uncertainty ranges (up - low). The CV was measured using population size estimated over previously known breeding locations. The blue gradient is based on the distribution quantile of the CV values with blue cell shading highlighting species with great discrepancies between the two population sizes estimated.**

Species	Untrimmed			Outlier-trimmed			CV (%)
	low	mean	up	low	mean	up	
Eurasian Blackcap	8.6M	9.5M	10.5M	8.6M	9.5M	10.5M	0.0
European Robin	7.7M	8.5M	9.5M	7.7M	8.5M	9.4M	2.3
Common Chaffinch	7.2M	8.0M	8.9M	7.2M	7.9M	8.8M	4.7
Eurasian Blue Tit	4.7M	5.4M	6.4M	4.7M	5.4M	6.3M	0.8
Great Tit	4.9M	5.4M	6.1M	4.9M	5.4M	6.1M	0.0
Eurasian Blackbird	4.0M	4.7M	5.6M	4.0M	4.5M	5.1M	24.1
European Goldfinch	3.0M	4.3M	6.1M	3.0M	4.1M	5.5M	16.3
Common Wood-Pigeon	3.2M	3.6M	4.0M	3.2M	3.6M	3.9M	1.5
Common Chiffchaff	3.1M	3.5M	4.0M	3.1M	3.5M	4.0M	1.0
Eurasian Wren	3.0M	3.5M	4.2M	3.0M	3.4M	3.9M	16.6
Barn Swallow	1.4M	2.1M	3.2M	1.4M	2.0M	2.9M	12.7
Eurasian Linnet	551.3K	1.7M	5.7M	551.3K	1.7M	5.6M	0.6
Eurasian Skylark	1.5M	2.3M	3.7M	1.5M	1.5M	2.1M	85.0
Short-toed Treecreeper	1.1M	1.3M	1.6M	1.1M	1.3M	1.6M	0.1
Firecrest	962.6K	1.3M	1.8M	962.6K	1.3M	1.8M	0.4
Dunnock	895.6K	1.2M	1.8M	895.6K	1.1M	1.5M	26.5

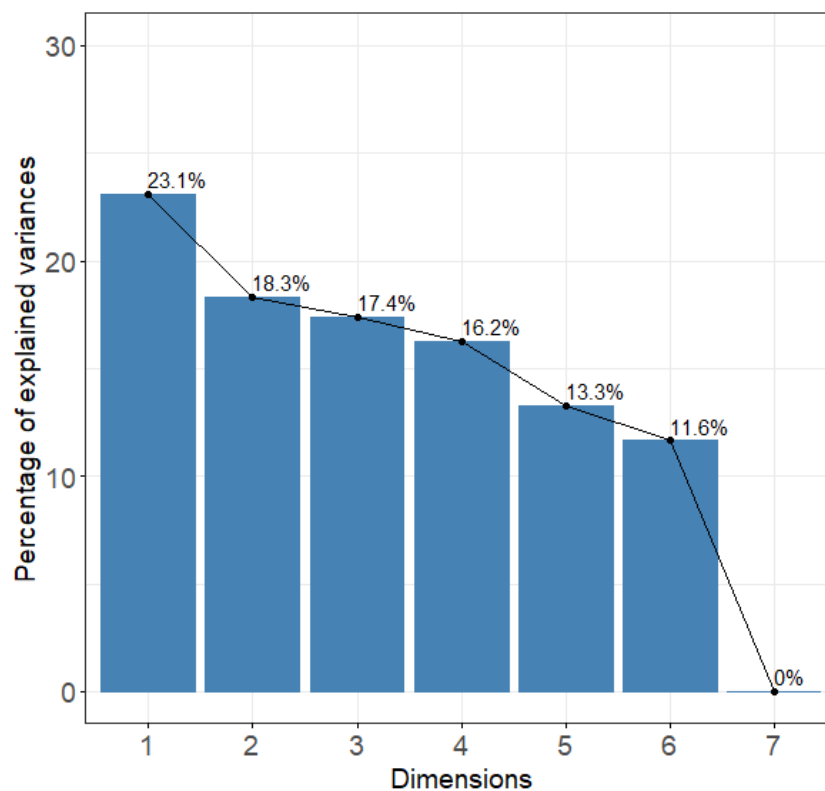
*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives*

*Supplementary information*

Song Thrush	966.4K	1.1M	1.3M	966.4K	1.1M	1.2M	9.7
White Wagtail	713.4K	1.0M	1.5M	713.4K	1.0M	1.5M	0.0
Eurasian Nuthatch	831.3K	1.0M	1.3M	831.3K	1.0M	1.3M	0.0
Goldcrest	1.3M	12.0M	197.5M	1.3M	980.6K	2.7M	139.5
Cirl Bunting	778.7K	1.2M	2.0M	778.7K	968.5K	1.4M	50.5
Black Redstart	717.1K	980.7K	1.3M	717.1K	966.4K	1.3M	4.8
Great Spotted Woodpecker	752.6K	976.6K	1.3M	752.6K	930.2K	1.2M	14.7
European Greenfinch	666.0K	886.9K	1.2M	666.0K	881.2K	1.2M	1.3
Common Nightingale	739.8K	885.5K	1.1M	739.8K	874.9K	1.0M	3.5
Hawfinch	422.3K	779.2K	1.6M	422.3K	747.4K	1.4M	16.7
European Stonechat	494.2K	734.9K	1.1M	494.2K	730.1K	1.1M	1.9
European Serin	532.4K	730.2K	1.0M	532.4K	714.8K	979.5K	4.2
Marsh Tit	440.1K	728.9K	1.3M	440.1K	672.8K	1.1M	21.6
Greater Whitethroat	475.9K	613.6K	795.6K	475.9K	613.5K	795.1K	0.1
Eurasian Sparrowhawk	30.2K	608.0K	7.6M	30.2K	603.5K	7.3M	2.9
Melodious Warbler	379.2K	500.9K	665.6K	379.2K	499.1K	662.2K	0.9
Crested Tit	312.2K	549.3K	1.2M	312.2K	492.5K	840.5K	33.5
Yellowhammer	370.4K	712.6K	1.8M	370.4K	453.6K	698.9K	87.3
Eurasian Bullfinch	269.1K	1.2M	12.4M	269.1K	437.2K	1.1M	123.7
Red-backed Shrike	244.1K	434.0K	818.4K	244.1K	419.9K	752.0K	8.7
Common Buzzard	270.6K	369.6K	516.2K	270.6K	367.8K	502.2K	4.1
European Turtle-Dove	265.3K	346.5K	454.3K	265.3K	344.1K	449.8K	1.7
Mistle Thrush	296.2K	373.0K	474.5K	296.2K	341.4K	417.2K	27.1
Western Bonelli's Warbler	148.8K	334.4K	852.3K	148.8K	322.5K	757.7K	10.2
Eurasian Golden Oriole	244.9K	318.0K	410.1K	244.9K	317.7K	409.3K	0.3
Eurasian Green Woodpecker	263.2K	316.2K	381.2K	263.2K	316.1K	380.8K	0.2
Ring-necked Pheasant	196.1K	294.5K	437.4K	196.1K	293.4K	434.3K	0.9
Western Yellow Wagtail	276.0K	651.1K	1.6M	276.0K	249.9K	410.5K	115.8
Tree Pipit	105.4K	197.0K	365.2K	105.4K	195.3K	355.6K	2.7
Corn Bunting	150.1K	196.2K	257.8K	150.1K	194.4K	254.3K	2.3
Wood Lark	153.1K	195.6K	249.9K	153.1K	192.9K	245.2K	3.5
Eurasian Kestrel	126.1K	192.5K	296.2K	126.1K	192.5K	295.9K	0.1
Common Cuckoo	129.8K	162.3K	200.0K	129.8K	162.3K	199.9K	0.1
Common Redstart	39.3K	129.8K	283.0K	39.3K	129.8K	283.0K	0.0
Willow Warbler	34.5K	113.5K	343.9K	34.5K	113.1K	341.6K	0.5
Black Woodpecker	26.6K	45.7K	79.5K	26.6K	44.0K	71.9K	11.0
Stock Dove	20.3K	54.9K	146.5K	20.3K	42.0K	86.6K	44.0
Eurasian Hoopoe	25.3K	41.6K	65.4K	25.3K	41.1K	64.5K	1.6

### **Selection of reduced PCA axes for habitat cover**

In order to limit the number of variables used to calibrate HDS models and help model convergence (Fig S3.1), we decided to restrain the number of PCA axes for habitat cover variables.



**Figure S3.2 - Variance explained scree plot from 7-class habitat cover PCA.**

Explained variance varied from 23.1% to 11.6% on retained six axes (Fig S3.2). After exploration of variable contributions and their associated map projections (Fig S3.3 - S3.4) we decided to use PCA axes 1, 2 and 5, corresponding to 54.71% of explained variance, in order to cover Forests, Urban, Pastures, Annual and Perennial crops gradients. Due to the scheme protocol not targeting wetland birds and habitats, we decided to discard the sixth PCA axis, which showed little to none variation across metropolitan France (Fig S3.4). This

Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives Supplementary information  
 specific habitat is monitored through other schemes such as the wetlands network (Moussy et al., 2022).

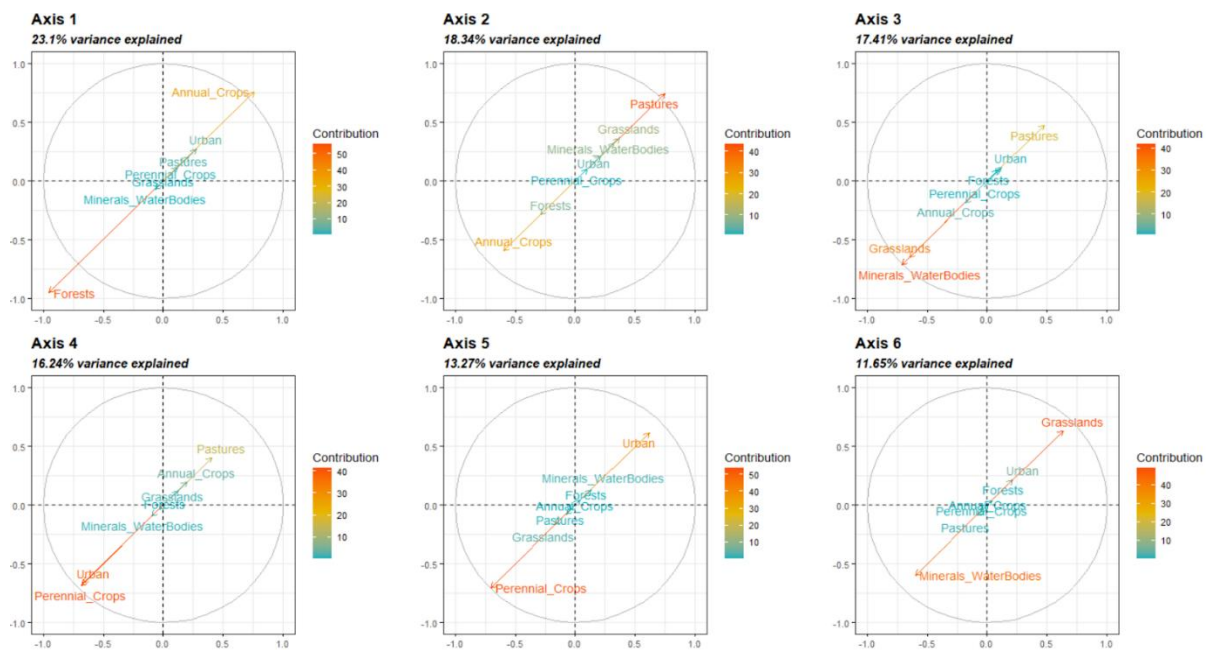


Figure S3.3 - Variable contribution plots for PCA axes. Each plot corresponds to the projection over the PCA axis in order to highlight the most influential variables.

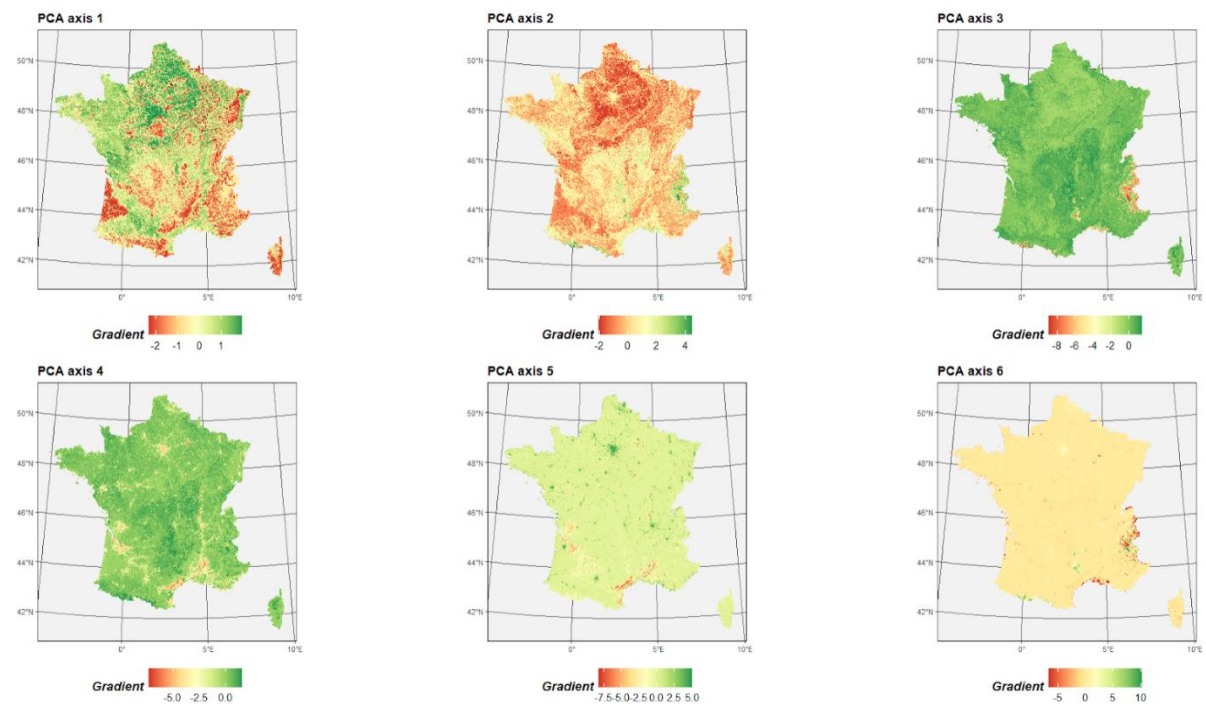


Figure S3.4 - Projection maps of PCA axes across metropolitan France.

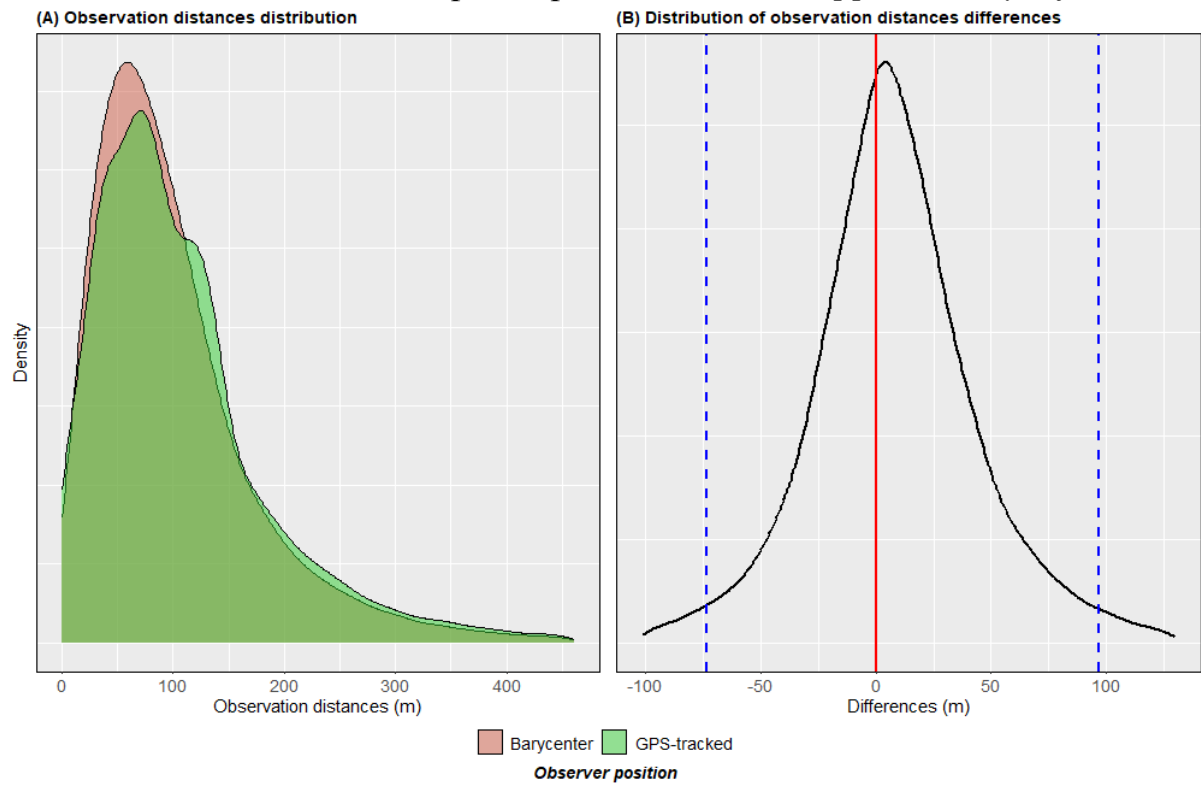
**Table S3.1 - Aggregation table for habitat cover using Theia OSO (Thierion et al., 2022) at 10m resolution.**

OSO class	Aggregation		
	Seven-class	Three-class	
1	Urban	Artificial	
2			
3			
4			
5	Annual crops	Open	
6			
7			
8			
9			
10			
11			
12			
14			Perennial crops
15			
18			Pastures
13			Grasslands
19			
20	Minerals / Water bodies		
21			
22			
23			
16	Forests	Forests	
17			

For our study, we gathered data using both precise observer locations and approximated observer locations based on the observations barycenter. In the latter case, the observer locations were approximated using observations barycenter. The goal of this appendix is to assess if this approximation could have an effect on a posteriori measured observation distance. We subsetting our initial dataset using 15 756 lists (*i.e.* 65% of the global dataset) with registered GPS location of observers. From this subset, we measured the barycenter of observations, in order to measure observation distance given the two types of observer positions, on the same observations. A similar pattern of observation distance distribution emerges between observer positions types (Fig S4.1A). When we estimate differences such as :

$$\Delta = Distances_{GPS} - Distances_{Barycenter}$$

We can see (Fig S4.1B) that the distribution of observation distance differences is slightly offset from 0, depicted as the solid red line. The average observation distance difference is 8.01 m. This suggests that on average observation distances measured using observers GPS locations are greater than those obtained using observation barycenters. This average difference should be taken into account, as the lowest possible resolution during the formation of distance bin classes, in order to limit misclassification of distance class while using these two types of distance data. For our study, the lowest resolution used for distance classes was for the Goldcrest (*Regulus regulus*) with 18m.



**Figure S4.1 - (A) Density plot of a posteriori measured observation distances given observer position being GPS-tracked or approximated as observations barycenter. (B) Distribution of differences between the two observers positions. Solid red line indicates 0 while dotted blue lines correspond to the 2.5% and 97.5% of the distribution.**

Hierarchical Distance sampling (*HDS*) seeks to model spatial variation of abundance while taking account of the detection process (Kéry & Royle, 2015). As Conventional Distance sampling assumes perfect detection (Buckland et al., 2015), *HDS* relaxes this assumption by assessing the probability that the individual is present and available for detection during survey occasions (Nichols et al., 2009) through multiple visits at the same site. Considering a population following Poisson distribution with mean  $\lambda$ , at each site  $i = 1, 2, \dots, I$  we have the local population  $M_i$  :

$$(i) M_i \sim \text{Poisson}(\lambda_i) \quad (1.1)$$

Given multiple visits  $j$  ( $j = 1, 2, \dots, J$ ), at site  $i$ , the number of individuals available for detection  $N_{i,j}$  follows binomial outcomes from the local population  $M_i$  with probability of availability  $\varphi_{i,j}$  :

$$(ii) N_{i,j} \sim \text{Binomial}(M_i, \varphi_{i,j}) \quad (1.2)$$

For each site  $i$  and visit  $j$ , observers measure the distance of observation between themselves and observed individuals. A vector of cell probabilities  $\pi_{i,j}$  derived from a detection function  $p$  (Buckland et al., 2015) assigns probabilities to distinct distance bins. Observation  $y_{i,j}$  can then be described as a multinomial outcome given the number of individuals available for detection and its distance:

$$(iii) y_{i,j} \sim \text{Multinomial}(N_{i,j}, \pi_{i,j}) \quad (1.3)$$

We used the secondary candidate set strategy (Fig S2.1) as described in Morin et al. (2020), where each state of the *HDS* is calibrated from primary candidate sets (Table S5.1)

while others are kept constant. We averaged predictions from the final candidate set in order to obtain prediction maps and population size estimates.

**Table S5.1 - Ensemble of sub-models tested in the secondary candidate set approach (Morin et al., 2020). Count column depicts the number of times each sub-processes was included in the final candidate sets (Figure X2 of the main text), across all studied species. (\*) For the hour effort covariate, we used minutes from sunrise estimated from site longitude, latitude and date of lists completion.**

<b>States</b>	<b>Sub-models</b>	<b>Counts</b>
Detection	~ Distance to roads	7
	~ Distance to roads + Proportion of artificial lands (100m)	0
	~ Distance to roads + Proportion of open lands (100m)	28
	~ Distance to roads + Proportion of forests (100m)	37
Availability	~ Julian date	26
	~ Julian date + Julian date <sup>2</sup>	33
	~ Hour (*)	16
	~ Hour + Hour <sup>2</sup>	19
	~ Julian date + Hour	26
	~ Julian date + Julian date <sup>2</sup> + Hour	38
	~ Julian date + Hour + Hour <sup>2</sup>	33
	~ Julian date + Julian date <sup>2</sup> + Hour + Hour <sup>2</sup>	43
Abundance	~ 3 Bioclimatics PCA axis + 3 Habitat cover PCA axis	63

### Comparison ArGeom and HDS

Table S6.1 - Results table of the 54 species with converging HDS model and their associated population sizes, in number of breeding pairs either estimated from the number of singing male encountered or the total number of adults divided by two, see table S2.1 for more details. Underlined-italic red species names (n = 8) correspond to species not used for the estimates comparison due to important variations between untrimmed and outlier-trimmed predictions (see, Table S2.3) not used for the estimates comparison.

Species habitat specialisation is derived from the FBBS (Godet et al., 2015). Species conservation status is derived from the IUCN red list for birds in France (UICN France et al., 2016). 2012 ArGeom estimates are retrieved from Issa and Muller (2015) population size estimates, while 2023 estimates correspond to their actualization using FBBS trends (Table S2.1). ArGeom uncertainty corresponds to the decimal logarithm of the difference between maximal and minimal estimated population. Shown HDS estimates correspond to the population size over the known breeding area of the 2012 atlas, see Table S6.2 for population size over the full metropolitan France.

Species mean density and availability estimated through HDS correspond to  $\lambda$  and  $\phi$  states intercepts in natural scales. Significant intercept ( $\alpha = 0.05$ ) are represented in bold, near significant intercept ( $\alpha = 0.1$ ) are followed with an (\*). Maximal distance of observation, in meters, represents the further distance bin after right-truncation of 5% of extreme distance values.

Additional informations in regards to HDS calibration are indicated, such as the chosen key function (hn : half-normal ; hr : hazard-rate); chosen distribution (P : Poisson; NB = Negative binomial) used for HDS abundance model, and measure of the overdispersion parameter ( $\hat{C}$ ).

The relative difference between HDS and ArGeom estimates ( $\delta_{mean}$ ) is represented over three categories : (i) ArGeom estimates greatly inferior to HDS estimates ( $\delta_{mean} < -0.2$ ) ; (ii) equivalent and (iii) ArGeom estimates greatly superior to HDS estimates ( $\delta_{mean} > 0.2$ ) of ArGeom population size estimations in red, beige and green cell shading respectively. ArGeom midpoints and HDS estimated population size are expressed in pairs.

Supplementary information

Species	Habitat specialization	Conservation status	ArGeom			HDS								$\hat{C}$	$\delta_{\text{mean}}$ (%)
			Midpoints		Range uncertainty	Estimation			Mean density (ind/km <sup>2</sup> )	Mean detection probability	Max distance (m)	Key function	Mixture		
			2012	2023		low	mean	up							
Eurasian Blackcap	Generalist	LC	6.5M	7.1M	6.51	8.6M	9.5M	10.5M	<b>16.21</b>	1.00	210	hr	P	1.01	-0.29
European Robin	Forest	LC	4.5M	7.2M	6.68	7.7M	8.5M	9.4M	<b>27.51</b>	<b>0.76</b>	170	hn	P	1.01	-0.17
Common Chaffinch	Generalist	LC	9.0M	8.3M	6.56	7.2M	7.9M	8.8M	<b>12.99</b>	1.00 <sup>(*)</sup>	225	hr	P	1.02	0.04
Eurasian Blue Tit	Generalist	LC	3.2M	3.5M	6.21	4.7M	5.4M	6.3M	<b>18.89</b>	<b>1.00</b>	140	hn	P	1.01	-0.42
Great Tit	Generalist	LC	5.5M	5.4M	6.47	4.9M	5.4M	6.1M	<b>17.69</b>	<b>0.99</b>	210	hn	P	1.04	0.00
Eurasian Blackbird	Generalist	LC	6.5M	6.1M	6.45	4.0M	4.5M	5.1M	<b>7.61</b>	<b>0.98</b>	290	hr	P	1.06	0.29
European Goldfinch	Urban	VU	1.5M	1.7M	6.06	3.0M	4.1M	5.5M	<b>12.13</b>	1.00	125	hr	P	1.02	-0.82
Common Wood-Pigeon	Generalist	LC	2.5M	3.7M	6.17	3.2M	3.6M	3.9M	<b>12.82</b>	1.00	295	hn	P	1.14	0.02
Common Chiffchaff	Forest	LC	3.2M	3.2M	6.17	3.1M	3.5M	4.0M	<b>12.24</b>	1.00 <sup>(*)</sup>	205	hn	P	1.00	-0.09
Eurasian Wren	Forest	LC	4.0M	3.4M	6.23	3.0M	3.4M	3.9M	<b>10.02</b>	1.00	180	hn	P	1.01	0.00
Barn Swallow	Urban	NT	1.3M	1.2M	5.89	1.4M	2.0M	2.9M	<b>5.23</b>	1.00	175	hr	P	1.20	-0.51
Eurasian Linnet	Farmland	VU	750.0K	1.0M	5.83	551.3K	1.7M	5.6M	2.58 <sup>(*)</sup>	0.84	140	hn	NB	1.17	-0.53
<i>Eurasian Skylark</i>	Farmland	NT	1.6M	1.5M	5.80	1.5M	1.5M	2.1M	<b>1.99</b>	1.00	275	hn	NB	1.25	-0.03
Short-toed Treecreeper	Forest	LC	1.2M	1.4M	5.90	1.1M	1.3M	1.6M	<b>4.66</b>	1.00 <sup>(*)</sup>	175	hn	P	1.04	0.05
Firecrest	Forest	LC	700.0K	809.2K	5.67	962.6K	1.3M	1.8M	<b>3.77</b>	1.00	115	hn	P	1.01	-0.46
Duncock	Generalist	LC	1.5M	1.1M	5.88	895.6K	1.1M	1.5M	<b>3.15</b>	<b>1.00</b>	170	hr	P	0.96	0.02
Song Thrush	Forest	LC	1.5M	1.6M	6.02	966.4K	1.1M	1.2M	<b>3.70</b>	<b>0.67</b>	330	hr	P	1.10	0.36
White Wagtail	/	LC	550.0K	523.0K	5.46	713.4K	1.0M	1.5M	<b>3.93</b>	0.77	175	hr	P	1.10	-0.66
Eurasian Nuthatch	Forest	LC	1.5M	1.0M	5.83	831.3K	1.0M	1.3M	<b>3.07</b>	0.45	240	hr	P	1.10	0.00
<i>Goldcrest</i>	Forest	NT	650.0K	350.4K	5.21	1.3M	980.6K	2.7M	<b>0.33</b>	1.00	90	hn	P	0.99	-0.95
<i>Cirl Bunting</i>	Farmland	LC	300.0K	344.1K	5.36	778.7K	968.5K	1.4M	<b>1.57</b>	<b>0.69</b>	220	hr	P	0.95	-0.95
Black Redstart	Urban	LC	950.0K	855.0K	5.80	717.1K	966.4K	1.3M	<b>1.30</b>	1.00	200	hr	P	1.06	-0.12

Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives

Supplementary information

Great Spotted Woodpecker	Forest	LC	550.0K	632.5K	5.54	752.6K	930.2K	1.2M	<b>3.04</b>	<b>0.59</b>	320	hn	NB	1.31	-0.38
European Greenfinch	Urban	VU	1.5M	862.5K	5.76	666.0K	881.2K	1.2M	<b>1.45</b>	<b>0.94</b>	215	hr	P	1.12	-0.02
Common Nightingale	Generalist	LC	750.0K	714.8K	5.68	739.8K	874.9K	1.0M	<b>2.19</b>	<b>1.00</b>	305	hn	P	1.12	-0.20
Hawfinch	Forest	LC	80.0K	66.0K	4.52	422.3K	747.4K	1.4M	<b>2.02</b>	0.34 <sup>(*)</sup>	130	hn	P	1.01	-1.68
European Stonechat	Farmland	NT	250.0K	307.8K	5.09	494.2K	730.1K	1.1M	1.23 <sup>(*)</sup>	1.00	160	hn	P	1.19	-0.81
European Serin	Urban	VU	375.0K	361.1K	5.38	532.4K	714.8K	979.5K	<b>1.62</b>	1.00	155	hn	P	1.02	-0.66
Marsh Tit	Forest	LC	800.0K	418.4K	5.32	440.1K	672.8K	1.1M	<b>1.49</b>	0.87	130	hr	P	0.99	-0.47
Greater Whitethroat	Farmland	LC	1.0M	944.0K	5.75	475.9K	613.5K	795.1K	<b>2.12</b>	1.00	185	hr	P	1.08	0.42
Eurasian Sparrowhawk	/	LC	47.3K	41.3K	3.87	30.2K	603.5K	7.3M	2.07	<b>0.05</b>	295	hn	P	1.13	-1.74
Melodious Warbler	Generalist	LC	675.0K	553.5K	5.57	379.2K	499.1K	662.2K	<b>1.73</b>	0.25	175	hr	P	1.08	0.10
<i>Crested Tit</i>	Forest	LC	450.0K	321.8K	5.33	312.2K	492.5K	840.5K	<b>0.29</b>	1.00	125	hn	P	0.96	-0.42
<i>Yellowhammer</i>	Farmland	VU	750.0K	379.5K	5.40	370.4K	453.6K	698.9K	<b>0.55</b>	1.00 <sup>(*)</sup>	230	hr	P	1.09	-0.18
<i>Eurasian Bullfinch</i>	Forest	VU	150.0K	78.0K	4.72	269.1K	437.2K	1.1M	0.63	0.32 <sup>(*)</sup>	120	hn	P	1.05	-1.39
Red-backed Shrike	Farmland	NT	150.0K	224.1K	5.17	244.1K	419.9K	752.0K	0.85	1.00	200	hn	NB	1.40	-0.61
Common Buzzard	Farmland	LC	160.0K	152.3K	4.28	270.6K	367.8K	502.2K	<b>1.37</b>	<b>0.37</b>	450	hr	P	1.01	-0.83
European Turtle-Dove	/	VU	400.0K	296.4K	5.17	265.3K	344.1K	449.8K	1.09	<b>1.00</b>	285	hr	P	1.07	-0.15
Mistle Thrush	Forest	LC	400.0K	386.4K	5.29	296.2K	341.4K	417.2K	<b>0.78</b>	1.00	410	hn	P	1.13	0.12
Western Bonelli's Warbler	Forest	LC	225.0K	241.0K	5.21	148.8K	322.5K	757.7K	<b>0.17</b>	<b>1.00</b>	135	hn	NB	1.45	-0.29
Eurasian Golden Oriole	Generalist	LC	150.0K	173.4K	5.06	244.9K	317.7K	409.3K	<b>0.58</b>	1.00	435	hr	P	1.05	-0.59
Eurasian Green Woodpecker	Generalist	LC	225.0K	207.5K	5.14	263.2K	316.1K	380.8K	<b>1.17</b>	<b>0.69</b>	500	hr	P	1.09	-0.42
Ring-necked Pheasant	/	LC	205.0K	201.3K	4.95	196.1K	293.4K	434.3K	<b>0.48</b>	<b>0.93</b>	380	hn	NB	1.18	-0.37
<i>Western Yellow Wagtail</i>	Farmland	LC	135.0K	116.5K	4.63	276.0K	249.9K	410.5K	<b>0.25</b>	1.00	190	hn	NB	1.25	-0.73
Tree Pipit	/	LC	375.0K	229.1K	5.18	105.4K	195.3K	355.6K	0.81	1.00	255	hr	NB	1.09	0.16
Corn Bunting	Farmland	LC	300.0K	363.0K	5.38	150.1K	194.4K	254.3K	<b>0.32</b>	<b>1.00</b>	270	hr	P	1.17	0.60
Wood Lark	Farmland	LC	140.0K	265.4K	5.06	153.1K	192.9K	245.2K	<b>0.82</b>	1.00 <sup>(*)</sup>	315	hn	P	1.16	0.32
Eurasian Kestrel	Farmland	NT	76.0K	81.5K	4.23	126.1K	192.5K	295.9K	<b>0.62</b>	0.94	310	hr	P	1.05	-0.81
Common Cuckoo	Generalist	LC	225.0K	208.3K	5.14	129.8K	162.3K	199.9K	<b>0.59</b>	<b>1.00</b>	645	hr	P	1.18	0.25

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives*

*Supplementary information*

Common Redstart	Urban	LC	120.0K	137.5K	4.84	39.3K	129.8K	283.0K	<b>0.33</b>	0.00	250	hr	NB	1.18	0.06	
Willow Warbler	Forest	NT	150.0K	83.7K	4.75	34.5K	113.1K	341.6K	<b>0.40</b>	0.92	130	hr	NB	1.36	-0.30	
Black Woodpecker	Forest	LC	32.5K	38.6K	4.25	26.6K	44.0K	71.9K	<b>0.09</b>	0.39	555	hn	P	1.02	-0.13	
<i>Stock Dove</i>	/	LC	45.0K	91.2K	4.78	20.3K	42.0K	86.6K	<b>0.10</b>	0.69	380	hn	P	0.91	0.74	
Eurasian Hoopoe	Farmland	LC	85.0K	77.9K	4.66	25.3K	41.1K	64.5K	<b>0.19</b>	1.00	455	hr	P	1.01	0.62	

**Estimation over metropolitan France and atlas grid filtering**

Table S6.2 - Table of the 54 studied species with converging HDS model and their associated population sizes. For the current study, we restricted the area for model prediction in order to correspond to previously known species' breeding locations from the last atlas (Issa & Muller, 2015). These estimates are depicted below the "Atlas grid" spanner.

The HDS model approach used in this paper can estimate population size over geographical areas not collected during the sampling process, while assessing possible issues in regards of extrapolation, see Table S2.3 and Fig 3.), allowing estimation of bird population size over metropolitan France, under the "Metropolitan France" spanner.

Population sizes are represented in the number of breeding pairs, for the HDS approach, this number is either estimated from the number of singing male encountered or the total number of adults divided by two, see table S2.1 for more details. Underlined-italic species names (n = 8) correspond to species not used for the estimates comparison due to important variations between untrimmed and outlier-trimmed population size estimates (see, Table S2.3).

Species	Atlas grid							Metropolitan France HDS (2021-2023)		
	ArGeom				HDS (2021-2023)					
	2012		2023					low	mean	up
	low	up	low	up						
Eurasian Blackcap	5.0M	8.0M	5.4M	8.7M	8.6M	9.5M	10.5M	8.8M	9.7M	10.7M
European Robin	3.0M	6.0M	4.8M	9.6M	7.7M	8.5M	9.4M	8.0M	8.9M	9.8M
Common Chaffinch	7.0M	11.0M	6.4M	10.1M	7.2M	7.9M	8.8M	7.4M	8.1M	9.0M
Eurasian Blue Tit	2.5M	4.0M	2.7M	4.4M	4.7M	5.4M	6.3M	4.9M	5.7M	6.6M
Great Tit	4.0M	7.0M	3.9M	6.9M	4.9M	5.4M	6.1M	5.0M	5.5M	6.2M
Eurasian Blackbird	5.0M	8.0M	4.7M	7.5M	4.0M	4.5M	5.1M	4.1M	4.6M	5.3M
European Goldfinch	1.0M	2.0M	1.1M	2.3M	3.0M	4.1M	5.5M	3.2M	4.3M	5.8M
Common Wood-Pigeon	2.0M	3.0M	2.9M	4.4M	3.2M	3.6M	3.9M	3.4M	3.7M	4.1M
Common Chiffchaff	2.5M	4.0M	2.5M	4.0M	3.1M	3.5M	4.0M	3.3M	3.7M	4.2M
Eurasian Wren	3.0M	5.0M	2.5M	4.2M	3.0M	3.4M	3.9M	3.2M	3.6M	4.1M

Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives

Supplementary information

Barn Swallow	900.0K	1.8M	779.4K	1.6M	1.4M	2.0M	2.9M	1.5M	2.1M	3.1M
Eurasian Linnet	500.0K	1.0M	673.5K	1.3M	551.3K	1.7M	5.6M	633.7K	2.0M	6.5M
<i>Eurasian Skylark</i>	1.3M	2.0M	1.2M	1.8M	1.5M	1.5M	2.1M	1.7M	1.7M	2.4M
Short-toed Treecreeper	900.0K	1.6M	1.0M	1.8M	1.1M	1.3M	1.6M	1.2M	1.4M	1.7M
Firecrest	500.0K	900.0K	578.0K	1.0M	962.6K	1.3M	1.8M	1.2M	1.6M	2.2M
Dunnock	1.0M	2.0M	759.0K	1.5M	895.6K	1.1M	1.5M	989.9K	1.2M	1.6M
Song Thrush	1.0M	2.0M	1.0M	2.1M	966.4K	1.1M	1.2M	1.0M	1.2M	1.3M
White Wagtail	400.0K	700.0K	380.4K	665.7K	713.4K	1.0M	1.5M	748.0K	1.1M	1.6M
Eurasian Nuthatch	1.0M	2.0M	682.0K	1.4M	831.3K	1.0M	1.3M	899.3K	1.1M	1.4M
<i>Goldcrest</i>	500.0K	800.0K	269.5K	431.2K	1.3M	980.6K	2.7M	2.1M	1.2M	3.4M
<i>Cirl Bunting</i>	200.0K	400.0K	229.4K	458.8K	778.7K	968.5K	1.4M	906.3K	1.1M	1.6M
Black Redstart	600.0K	1.3M	540.0K	1.2M	717.1K	966.4K	1.3M	742.6K	1.0M	1.4M
Great Spotted Woodpecker	400.0K	700.0K	460.0K	805.0K	752.6K	930.2K	1.2M	794.9K	987.9K	1.3M
European Greenfinch	1.0M	2.0M	575.0K	1.1M	666.0K	881.2K	1.2M	693.2K	919.8K	1.2M
Common Nightingale	500.0K	1.0M	476.5K	953.0K	739.8K	874.9K	1.0M	806.9K	955.9K	1.1M
Hawfinch	60.0K	100.0K	49.5K	82.5K	422.3K	747.4K	1.4M	422.3K	747.4K	1.4M
European Stonechat	200.0K	300.0K	246.2K	369.3K	494.2K	730.1K	1.1M	558.2K	832.9K	1.2M
European Serin	250.0K	500.0K	240.8K	481.5K	532.4K	714.8K	979.5K	589.8K	792.9K	1.1M
Marsh Tit	600.0K	1.0M	313.8K	523.0K	440.1K	672.8K	1.1M	553.6K	847.9K	1.4M
Greater Whitethroat	700.0K	1.3M	660.8K	1.2M	475.9K	613.5K	795.1K	547.7K	714.4K	938.7K
Eurasian Sparrowhawk	43.1K	51.5K	37.7K	45.0K	30.2K	603.5K	7.3M	37.0K	725.2K	8.8M
Melodious Warbler	450.0K	900.0K	369.0K	738.0K	379.2K	499.1K	662.2K	425.6K	564.6K	755.5K
<i>Crested Tit</i>	300.0K	600.0K	214.5K	429.0K	312.2K	492.5K	840.5K	372.8K	591.7K	1.0M
<i>Yellowhammer</i>	500.0K	1.0M	253.0K	506.0K	370.4K	453.6K	698.9K	429.5K	494.5K	771.7K
<i>Eurasian Bullfinch</i>	100.0K	200.0K	52.0K	104.0K	269.1K	437.2K	1.1M	355.1K	550.5K	1.4M
Red-backed Shrike	100.0K	200.0K	149.4K	298.8K	244.1K	419.9K	752.0K	244.1K	419.9K	752.0K
Common Buzzard	150.0K	170.0K	142.8K	161.8K	270.6K	367.8K	502.2K	281.0K	383.2K	523.7K

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives*

*Supplementary information*

European Turtle-Dove	300.0K	500.0K	222.3K	370.5K	265.3K	344.1K	449.8K	295.2K	383.6K	504.4K
Mistle Thrush	300.0K	500.0K	289.8K	483.0K	296.2K	341.4K	417.2K	322.5K	369.9K	453.6K
Western Bonelli's Warbler	150.0K	300.0K	160.7K	321.3K	148.8K	322.5K	757.7K	210.2K	464.5K	1.1M
Eurasian Golden Oriole	100.0K	200.0K	115.6K	231.2K	244.9K	317.7K	409.3K	286.2K	377.9K	495.7K
Eurasian Green Woodpecker	150.0K	300.0K	138.3K	276.6K	263.2K	316.1K	380.8K	274.5K	330.3K	398.7K
Ring-necked Pheasant	160.0K	250.0K	157.1K	245.5K	196.1K	293.4K	434.3K	252.7K	384.8K	579.8K
<i>Western Yellow Wagtail</i>	110.0K	160.0K	94.9K	138.1K	276.0K	249.9K	410.5K	353.3K	385.5K	715.0K
Tree Pipit	250.0K	500.0K	152.8K	305.5K	105.4K	195.3K	355.6K	131.2K	249.4K	467.0K
Corn Bunting	200.0K	400.0K	242.0K	484.0K	150.1K	194.4K	254.3K	184.1K	240.9K	318.3K
Wood Lark	110.0K	170.0K	208.6K	322.3K	153.1K	192.9K	245.2K	218.5K	281.0K	363.0K
Eurasian Kestrel	68.0K	84.0K	73.0K	90.1K	126.1K	192.5K	295.9K	130.7K	200.0K	308.5K
Common Cuckoo	150.0K	300.0K	138.9K	277.8K	129.8K	162.3K	199.9K	138.8K	174.2K	215.1K
Common Redstart	90.0K	150.0K	103.1K	171.9K	39.3K	129.8K	283.0K	47.1K	171.6K	405.7K
Willow Warbler	100.0K	200.0K	55.8K	111.6K	34.5K	113.1K	341.6K	49.4K	178.1K	604.0K
Black Woodpecker	25.0K	40.0K	29.7K	47.5K	26.6K	44.0K	71.9K	29.0K	49.3K	82.2K
<i>Stock Dove</i>	30.0K	60.0K	60.8K	121.6K	20.3K	42.0K	86.6K	307.4K	78.1K	173.3K
Eurasian Hoopoe	60.0K	110.0K	55.0K	100.9K	25.3K	41.1K	64.5K	27.4K	46.6K	78.2K

*HDS species parameter*

Table S6.3 - Results table of the 54 species with converging HDS model and their associated population sizes, in number of breeding pairs either estimated from the number of singing male encountered or the total number of adults divided by two, see table S2.1 for more details.

Underlined-italic red species names (n = 8) correspond to species not used for the estimates comparison due to important variations between untrimmed and outlier-trimmed treatment (see, Table S2.3) not used for the estimates comparison.

Species mean density and availability estimated through HDS correspond to  $\lambda$  and  $\varphi$  states intercepts in natural scales. Species detectability ( $\sigma$ ) and scale parameters, for species using hazard-rate detection functions are on the log scale, with higher values corresponding to higher probability of detecting individuals at a given distance of observation. Significant estimated parameters ( $\alpha = 0.05$ ) are represented in bold, near significant intercept ( $\alpha = 0.1$ ) are followed with an (\*).

Additional informations in regards to HDS calibration are indicated, such as the chosen key function (hn : half-normal ; hr : hazard-rate); chosen distribution (P : Poisson; NB = Negative binomial) used for HDS abundance model, and measure of the overdispersion parameter ( $\hat{C}$ ).

Species	Mean density (ind/km <sup>2</sup> )	Mean detection probability	Key function $\sigma$ parameter	Scale parameter	Key function	Mixture	$\hat{C}$
Eurasian Blackcap	<b>16.21</b>	1.00	<b>4.60</b>	<b>1.57</b>	hr	P	1.01
European Robin	<b>27.51</b>	<b>0.76</b>	<b>4.03</b>	/	hn	P	1.01
Common Chaffinch	<b>12.99</b>	1.00 <sup>(*)</sup>	<b>4.66</b>	<b>1.56</b>	hr	P	1.02
Eurasian Blue Tit	<b>18.89</b>	<b>1.00</b>	<b>3.85</b>	/	hn	P	1.01
Great Tit	<b>17.69</b>	<b>0.99</b>	<b>4.19</b>	/	hn	P	1.04
Eurasian Blackbird	<b>7.61</b>	<b>0.98</b>	<b>4.82</b>	<b>1.56</b>	hr	P	1.06
European Goldfinch	<b>12.13</b>	1.00	<b>3.35</b>	<b>0.91</b>	hr	P	1.02
Common Wood-Pigeon	<b>12.82</b>	1.00	<b>4.57</b>	/	hn	P	1.14
Common Chiffchaff	<b>12.24</b>	1.00 <sup>(*)</sup>	<b>4.28</b>	/	hn	P	1.00
Eurasian Wren	<b>10.02</b>	1.00	<b>4.20</b>	/	hn	P	1.01
Barn Swallow	<b>5.23</b>	1.00	<b>3.75</b>	<b>1.00</b>	hr	P	1.20
Eurasian Linnet	2.58 <sup>(*)</sup>	0.84	<b>3.78</b>	/	hn	NB	1.17
<i>Eurasian Skylark</i>	<b>1.99</b>	1.00	<b>4.63</b>	/	hn	NB	1.25
Short-toed Treecreeper	<b>4.66</b>	1.00 <sup>(*)</sup>	<b>4.05</b>	/	hn	P	1.04
Firecrest	<b>3.77</b>	1.00	<b>3.73</b>	/	hn	P	1.01
Dunnock	<b>3.15</b>	<b>1.00</b>	<b>3.98</b>	<b>1.17</b>	hr	P	0.96
Song Thrush	<b>3.70</b>	<b>0.67</b>	<b>4.91</b>	<b>1.46</b>	hr	P	1.10
White Wagtail	<b>3.93</b>	0.77	<b>3.66</b>	<b>1.03</b>	hr	P	1.10
Eurasian Nuthatch	<b>3.07</b>	0.45	<b>4.42</b>	<b>1.32</b>	hr	P	1.10
<i>Goldcrest</i>	<b>0.33</b>	1.00	<b>3.55</b>	/	hn	P	0.99
<i>Cirl Bunting</i>	<b>1.57</b>	<b>0.69</b>	<b>4.72</b>	<b>1.75</b>	hr	P	0.95
Black Redstart	<b>1.30</b>	1.00	<b>4.35</b>	<b>1.40</b>	hr	P	1.06
Great Spotted Woodpecker	<b>3.04</b>	<b>0.59</b>	<b>4.54</b>	/	hn	NB	1.31

*Chapitre 1 : Estimation des tailles de population d'oiseaux communs nicheurs à l'aide de données de sciences participatives*

*Supplementary information*

European Greenfinch	<b>1.45</b>	<b>0.94</b>	<b>4.43</b>	<b>1.47</b>	hr	P	1.12
Common Nightingale	<b>2.19</b>	<b>1.00</b>	<b>4.62</b>	/	hn	P	1.12
Hawfinch	<b>2.02</b>	0.34 <sup>(*)</sup>	<b>3.79</b>	/	hn	P	1.01
European Stonechat	1.23 <sup>(*)</sup>	1.00	<b>4.10</b>	/	hn	P	1.19
European Serin	<b>1.62</b>	1.00	<b>3.97</b>	/	hn	P	1.02
Marsh Tit	<b>1.49</b>	0.87	<b>3.58</b>	<b>1.10</b>	hr	P	0.99
Greater Whitethroat	<b>2.12</b>	1.00	<b>4.46</b>	<b>1.52</b>	hr	P	1.08
Eurasian Sparrowhawk	2.07	<b>0.05</b>	<b>4.40</b>	/	hn	P	1.13
Melodious Warbler	<b>1.73</b>	0.25	<b>4.55</b>	<b>1.56</b>	hr	P	1.08
<i>Crested Tit</i>	<b>0.29</b>	1.00	<b>3.77</b>	/	hn	P	0.96
<i>Yellowhammer</i>	<b>0.55</b>	1.00 <sup>(*)</sup>	<b>4.56</b>	<b>1.34</b>	hr	P	1.09
<i>Eurasian Bullfinch</i>	0.63	0.32 <sup>(*)</sup>	<b>3.76</b>	/	hn	P	1.05
Red-backed Shrike	0.85	1.00	<b>4.11</b>	/	hn	NB	1.40
Common Buzzard	<b>1.37</b>	<b>0.37</b>	<b>4.96</b>	<b>1.43</b>	hr	P	1.01
European Turtle-Dove	1.09	<b>1.00</b>	<b>4.95</b>	<b>1.65</b>	hr	P	1.07
Mistle Thrush	<b>0.78</b>	1.00	<b>4.91</b>	/	hn	P	1.13
Western Bonelli's Warbler	<b>0.17</b>	<b>1.00</b>	<b>4.20</b>	/	hn	NB	1.45
Eurasian Golden Oriole	<b>0.58</b>	1.00	<b>5.14</b>	<b>1.55</b>	hr	P	1.05
Eurasian Green Woodpecker	<b>1.17</b>	<b>0.69</b>	<b>5.22</b>	<b>1.49</b>	hr	P	1.09
Ring-necked Pheasant	<b>0.48</b>	<b>0.93</b>	<b>4.81</b>	/	hn	NB	1.18
<i>Western Yellow Wagtail</i>	<b>0.25</b>	1.00	<b>4.40</b>	/	hn	NB	1.25
Tree Pipit	0.81	1.00	<b>4.68</b>	<b>1.53</b>	hr	NB	1.09
Corn Bunting	<b>0.32</b>	<b>1.00</b>	<b>4.78</b>	<b>1.53</b>	hr	P	1.17
Wood Lark	<b>0.82</b>	1.00 <sup>(*)</sup>	<b>4.76</b>	/	hn	P	1.16
Eurasian Kestrel	<b>0.62</b>	0.94	<b>4.51</b>	<b>1.05</b>	hr	P	1.05
Common Cuckoo	<b>0.59</b>	<b>1.00</b>	<b>5.46</b>	<b>1.40</b>	hr	P	1.18
Common Redstart	<b>0.33</b>	0.00	<b>4.71</b>	<b>1.53</b>	hr	NB	1.18
Willow Warbler	<b>0.40</b>	0.92	<b>4.30</b>	<b>1.52</b>	hr	NB	1.36
Black Woodpecker	<b>0.09</b>	0.39	<b>5.15</b>	/	hn	P	1.02
<i>Stock Dove</i>	<b>0.10</b>	0.69	<b>4.89</b>	/	hn	P	0.91
Eurasian Hoopoe	<b>0.19</b>	1.00	<b>5.14</b>	<b>1.41</b>	hr	P	1.01

***Comparison ArGeom and HDS estimates with German population size***

Table S6.4 - Comparison of population size estimates from ArGeom and HDS approach, see more detail in table S6.1, with German population size extracted from the European Red List of Birds (BirdLife International, 2021). Estimated German population size relied on point count and territory mapping methods (Gedeon et al., 2015). For comparable estimates between the two countries, we applied a coefficient computed on the ratio of countries area (~1.54).

Species	Scientific names	German population sizes		German population sizes (area adjusted)		ArGeom midpoint 2023	HDS mean	$\delta_{\text{mean}}$
		low	high	low	high			
Eurasian Blackcap	<i>Sylvia atricapilla</i>	4 650 000	6 150 000	7 174 047	9 488 255	7 072 000	9 495 742	-0,29
Firecrest	<i>Regulus ignicapilla</i>	1 250 000	1 850 000	1 928 507	2 854 191	809 200	1 290 838	-0,46
Eurasian Blue tit	<i>Cyanistes caeruleus</i>	3 250 000	4 800 000	5 014 119	7 405 468	3 545 750	5 441 758	-0,42
Common Cuckoo	<i>Cuculus canorus</i>	38 000	62 000	58 627	95 654	208 350	162 302	0,25
Corn Bunting	<i>Emberiza calandra</i>	16 500	29 000	25 456	44 741	363 000	194 425	0,6
Common Whitethroat	<i>Curruca communis</i>	600 000	950 000	925 683	1 465 665	944 000	613 495	0,42

S7: Complementary analysis of the comparison (in response to reviewers)

For this complementary analysis, we computed mean using ArGeom and HDS upper bounds, instead of the midpoint and mean used in the main text (see fig 5 of the main text). After calibration of the same model described in the “Variation of estimated population sizes between the two approaches” section in the main text, we observe the same response pattern of mean across species detection probabilities and ArGeom estimated range uncertainty than those obtained in the main text. Compare to figure 7, in the main text, despite keeping the significant positive effect of ArGeom uncertainty to mean variation, we see that, on average (*i.e.* in terms of average detection probabilities and ArGeom uncertainty from the studied species), there is no significant differences between ArGeom and HDS upper bounds.

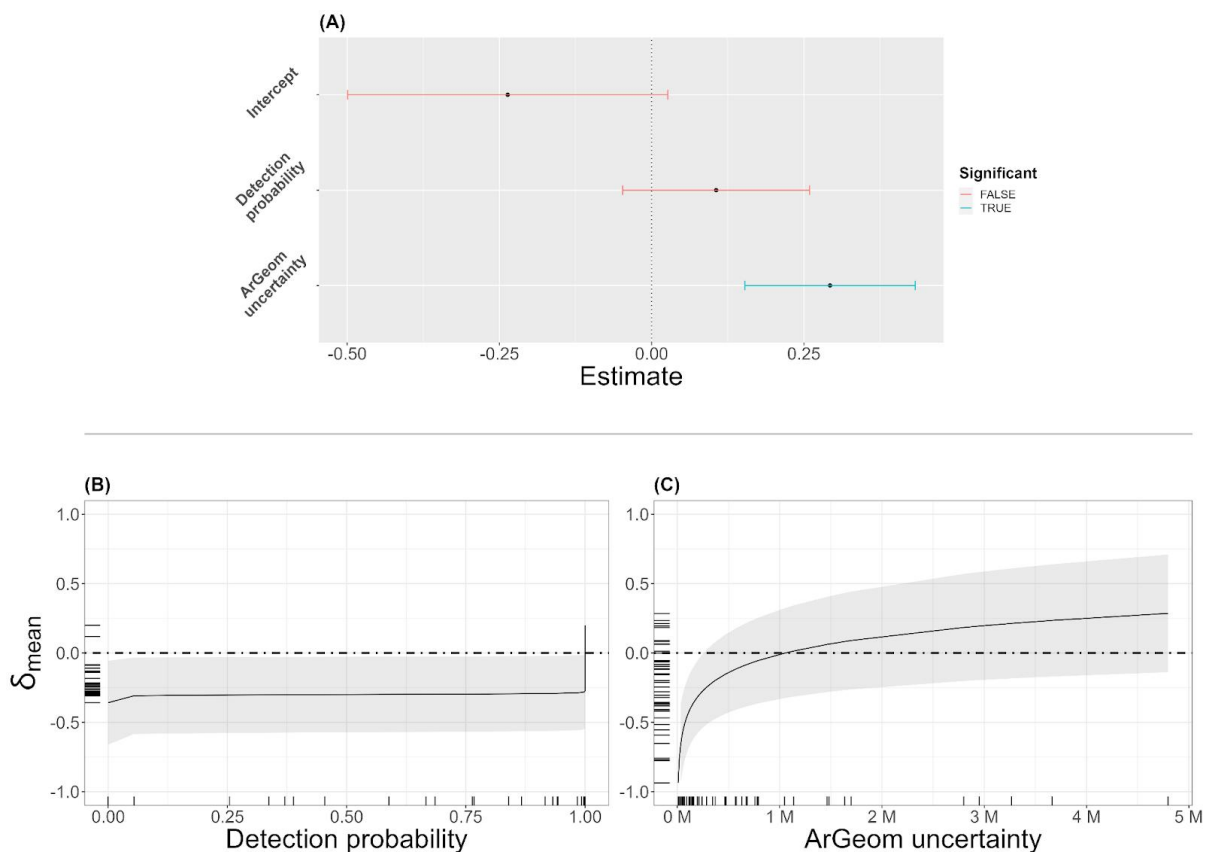


Figure S7.1 - Results from the PGLMM. (A) Confidence intervals of the model coefficient, parameters, coefficients significantly different from 0 are represented in blue. Marginal effect plots of population size estimates differences ( $\delta_{mean}$ ).  $\delta_{mean}$  responses are predicted over gradients of focal terms, either species detection probabilities (B) or ArGeom reported uncertainties (C), while other covariates are held constant at their mean. Species detection probabilities and ArGeom uncertainty are represented on their natural scales, after inverse logit and inverse decimal logarithm transformation, respectively. Dot-dash line corresponds to a  $\delta_{mean}$  of 0, signifying estimated population sizes convergence by the two approaches, negative and positive values of  $\delta_{mean}$  reflect lower and higher population size estimates of ArGeom relative to those obtained using HDS.

### III.D References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Barker RJ, Schofield MR, Link WA, Sauer JR (2018) On the Reliability of N-mixture Models for Count Data. *Biometrics*, **74**, 369–377. <https://doi.org/10.1111/biom.12734>
- BirdLife International (2021) European Red List of Birds. *Office of the European Union*.
- Bouchet PJ, Miller DL, Roberts JJ, Mannocci L, Harris CM, Thomas L (2020) dsmextra: Extrapolation assessment tools for density surface models. *Methods in Ecology and Evolution*, **11**, 1464–1469. <https://doi.org/10.1111/2041-210X.13469>
- Boulinier T, Nichols JD, Sauer JR, Hines JE, Pollock KH (1998) Estimating Species Richness: The Importance of Heterogeneity in Species Detectability. *Ecology*, **79**, 1018–1028. [https://doi.org/10.1890/0012-9658\(1998\)079\[1018:ESRTIO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1998)079[1018:ESRTIO]2.0.CO;2)
- Brlík V, Šilarová E, Škorpilová J, Alonso H, Anton M, Aunins A, Benkő Z, Biver G, Busch M, Chodkiewicz T, Chylarecki P, Coombes D, de Carli E, del Moral JC, Derouaux A, Escandell V, Eskildsen DP, Fontaine B, Foppen RPB, Gamero A, Gregory RD, Harris S, Herrando S, Hristov I, Husby M, Ieronymidou C, Jiquet F, Kålås JA, Kamp J, Kmecl P, Kurlavičius P, Lehtikainen A, Lewis L, Lindström Å, Manolopoulos A, Martí D, Massimino D, Moshøj C, Nellis R, Noble D, Paquet A, Paquet J-Y, Portolou D, Ramírez I, Redel C, Reif J, Ridzoň J, Schmid H, Seaman B, Silva L, Soldaat L, Spasov S, Staneva A, Szép T, Florenzano GT, Teufelbauer N, Trautmann S, van der Meij T, van Strien A, van Turnhout C, Vermeersch G, Vermouzek Z, Vikstrøm T, Voříšek P, Weiserbs A, Klvaňová A (2021) Long-term and large-scale multispecies dataset tracking population changes of common European breeding birds. *Scientific Data*, **8**, 21. <https://doi.org/10.1038/s41597-021-00804-2>
- Buckland ST, Anderson DR, Burnham KP, Laake J, Borchers D, Thomas L (2001) *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford ; New York.
- Buckland ST, Johnston A (2017) Monitoring the biodiversity of regions: Key principles and possible pitfalls. *Biological Conservation*, **214**, 23–34. <https://doi.org/10.1016/j.biocon.2017.07.034>
- Buckland ST, Rexstad EA, Marques TA, Oedekoven CS (2015) *Distance Sampling: Methods and Applications*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-19219-2>
- Burleigh JG, Kimball RT, Braun EL (2015) Building the avian tree of life using a large-scale, sparse supermatrix. *Molecular Phylogenetics and Evolution*, **84**, 53–63. <https://doi.org/10.1016/j.ympev.2014.12.003>
- Callaghan CT, Nakagawa S, Cornwell WK (2021) Global abundance estimates for 9,700 bird species. *Proceedings of the National Academy of Sciences*, **118**. <https://doi.org/10.1073/pnas.2023170118>
- Callaghan CT, Nakagawa S, Cornwell WK (2022) Reply to Robinson et al.: Data integration will form the basis of future abundance estimates. *Proceedings of the National Academy of Sciences*, **119**, e2117920119. <https://doi.org/10.1073/pnas.2117920119>

- Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM (2015) Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, **1**, e1400253. <https://doi.org/10.1126/sciadv.1400253>
- Chandler RB, Royle JA, King DI (2011) Inference about density and temporary emigration in unmarked populations. *Ecology*, **92**, 1429–1435. <https://doi.org/10.1890/10-2433.1>
- Chandler M, See L, Copas K, Bonde AMZ, López BC, Danielsen F, Legind JK, Masinde S, Miller-Rushing AJ, Newman G, Rosemartin A, Turak E (2017) Contribution of citizen science towards international biodiversity monitoring. *Biological Conservation*, **213**, 280–294. <https://doi.org/10.1016/j.biocon.2016.09.004>
- Clavel J, Julliard R, Devictor V (2011) Worldwide decline of specialist species: toward a global functional homogenization? *Frontiers in Ecology and the Environment*, **9**, 222–228. <https://doi.org/10.1890/080216>
- Clement MJ, Converse SJ, Royle JA (2017) Accounting for imperfect detection of groups and individuals when estimating abundance. *Ecology and Evolution*, **7**, 7304–7310. <https://doi.org/10.1002/ece3.3284>
- Comolet-Tirman J, Quaintenne G, Sibley J-P, Wroza S, Bal G, Champagnon J, Couzi L, Czajkowski M-A, Denis P, Frochot B, Muller Y (2022) Estimation des populations d'oiseaux nicheurs de France métropolitaine. *Alauda*, **90**, 133–150.
- Connors BM, Cooper AB, Peterman RM, Dulvy NK (2014) The false classification of extinction risk in noisy environments. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20132935. <https://doi.org/10.1098/rspb.2013.2935>
- Cote C, Troncon C, Troncon C, Troncon C (2021) ROUTE 500® Version 3.0 - Descriptif de contenu. , 27.
- Couvet D, Ducarme F (2014) Reconciliation ecology, from biological to social challenges. *Revue d'ethnoécologie*. <https://doi.org/10.4000/ethnoecologie.1979>
- Cruz-Cárdenas G, López-Mata L, Villaseñor JL, Ortiz E (2014) Potential species distribution modeling and the use of principal component analysis as predictor variables. *Revista Mexicana de Biodiversidad*, **85**, 189–199. <https://doi.org/10.7550/rmb.36723>
- Devictor V, van Swaay C, Brereton T, Brotons L, Chamberlain D, Heliölä J, Herrando S, Julliard R, Kuussaari M, Lindström Å, Reif J, Roy DB, Schweiger O, Settele J, Stefanescu C, Van Strien A, Van Turnhout C, Vermouzek Z, WallisDeVries M, Wynhoff I, Jiguet F (2012) Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change*, **2**, 121–124. <https://doi.org/10.1038/nclimate1347>
- Devictor V, Whittaker RJ, Beltrame C (2010) Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions*, **16**, 354–362. <https://doi.org/10.1111/j.1472-4642.2009.00615.x>
- Donald PF, Fuller RJ (1998) Ornithological atlas data: a review of uses and limitations. *Bird Study*, **45**, 129–145. <https://doi.org/10.1080/00063659809461086>
- Duarte A, Adams MJ, Peterson JT (2018) Fitting N-mixture models to count data with unmodeled heterogeneity: Bias, diagnostics, and alternative approaches. *Ecological Modelling*, **374**, 51–59. <https://doi.org/10.1016/j.ecolmodel.2018.02.007>

- Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Farrier D, Whelan R, Mooney C (2007) Threatened species listing as a trigger for conservation action. *Environmental Science & Policy*, **10**, 219–229. <https://doi.org/10.1016/j.envsci.2006.12.001>
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, **37**, 4302–4315. <https://doi.org/10.1002/joc.5086>
- Fiske I, Chandler R (2011) unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance. *Journal of Statistical Software*, **43**, 1–23. <https://doi.org/10.18637/jss.v043.i10>
- Fithian W, Elith J, Hastie T, Keith DA (2015) Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in Ecology and Evolution*, **6**, 424–438. <https://doi.org/10.1111/2041-210X.12242>
- Fletcher RJ, Hefley TJ, Robertson EP, Zuckerberg B, McCleery RA, Dorazio RM (2019) A practical guide for combining data to model species distributions. *Ecology*, **100**, e02710. <https://doi.org/10.1002/ecy.2710>
- Fontaine A, Simard A, Brunet N, Elliott KH (2022) Scientific contributions of citizen science applied to rare or threatened animals. *Conservation Biology*, **36**, e13976. <https://doi.org/10.1111/cobi.13976>
- Freckleton RP (2020) Conservation decisions in the face of uncertainty. In: *Conservation Research, Policy and Practice Ecological Reviews*. (eds Vickery JA, Ockendon N, Pettorelli N, Brotherton PNM, Sutherland WJ, Davies ZG), pp. 183–195. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781108638210.011>
- Fuller RJ, Langslow DR (1984) Estimating numbers of birds by point counts: how long should counts last? *Bird Study*, **31**, 195–202. <https://doi.org/10.1080/00063658409476841>
- Gaston KJ, Fuller RA (2008) Commonness, population depletion and conservation biology. *Trends in Ecology & Evolution*, **23**, 14–19. <https://doi.org/10.1016/j.tree.2007.11.001>
- Gaüzère P, Barbaro L, Calatayud F, Princé K, Devictor V, Raison L, Sirami C, Balent G (2020) Long-term effects of combined land-use and climate changes on local bird communities in mosaic agricultural landscapes. *Agriculture, Ecosystems & Environment*, **289**, 106722. <https://doi.org/10.1016/j.agee.2019.106722>
- Gaüzère P, Devictor V (2021) Mismatches between birds' spatial and temporal dynamics reflect their delayed response to global changes. *Oikos*, **130**, 1284–1296. <https://doi.org/10.1111/oik.08289>
- Gaüzère P, Doucier G, Devictor V, Kéfi S (2019) A framework for estimating species-specific contributions to community indicators. *Ecological Indicators*, **99**, 74–82. <https://doi.org/10.1016/j.ecolind.2018.11.069>
- Gedeon K, Grüneberg C, Mitschke A, Sudfeldt C, Eickhorst W, Fischer S, Flade M, Frick S, Geiersberger I, Koop B, Kramer M, Krüger T, Roth N, Ryslavy T, Stübing S (2015) *Atlas Deutscher Brutvogelarten: Atlas of German Breeding Birds*. Dachverband Deutscher Avifaunisten, Münster, Westf.

- Gelfand AE, Kim H-J, Sirmans CF, Banerjee S (2003) Spatial Modeling With Spatially Varying Coefficient Processes. *Journal of the American Statistical Association*, **98**, 387–396. <https://doi.org/10.1198/016214503000170>
- Germain RR, Feng S, Chen G, Graves GR, Tobias JA, Rahbek C, Lei F, Fjeldså J, Hosner PA, Gilbert MTP, Zhang G, Nogués-Bravo D (2023) Species-specific traits mediate avian demographic responses under past climate change. *Nature Ecology & Evolution*, **7**, 862–872. <https://doi.org/10.1038/s41559-023-02055-3>
- Godet L, Gaüzere P, Jiguet F, Devictor V (2015) Dissociating several forms of commonness in birds sheds new light on biotic homogenization. *Global Ecology and Biogeography*, **24**, 416–426. <https://doi.org/10.1111/geb.12266>
- Guillera-Arroita G, Lahoz-Monfort JJ, MacKenzie DI, Wintle BA, McCarthy MA (2014) Ignoring Imperfect Detection in Biological Surveys Is Dangerous: A Response to ‘Fitting and Interpreting Occupancy Models’. *PLOS ONE*, **9**, e99571. <https://doi.org/10.1371/journal.pone.0099571>
- Hong P, Li Z, Yang Q, Deng W, Xu Y, Tobias JA, Wang S (2023) Functional traits and environment jointly determine the spatial scaling of population stability in North American birds. *Ecology*, **104**, e3973. <https://doi.org/10.1002/ecy.3973>
- Hurlbert AH, Liang Z (2012) Spatiotemporal Variation in Avian Migration Phenology: Citizen Science Reveals Effects of Climate Change. *PLOS ONE*, **7**, e31662. <https://doi.org/10.1371/journal.pone.0031662>
- Inger R, Gregory R, Duffy JP, Stott I, Voříšek P, Gaston KJ (2015) Common European birds are declining rapidly while less abundant species’ numbers are rising. *Ecology Letters*, **18**, 28–36. <https://doi.org/10.1111/ele.12387>
- Ingram DJ, Ferreira GB, Jones KE, Mace GM (2021) Targeting Conservation Actions at Species Threat Response Thresholds. *Trends in Ecology & Evolution*, **36**, 216–226. <https://doi.org/10.1016/j.tree.2020.11.004>
- Isaac NJB, Jarzyna MA, Keil P, Dambly LI, Boersch-Supan PH, Browning E, Freeman SN, Golding N, Guillera-Arroita G, Henrys PA, Jarvis S, Lahoz-Monfort J, Pagel J, Pescott OL, Schmucki R, Simmonds EG, O’Hara RB (2020) Data Integration for Large-Scale Models of Species Distributions. *Trends in Ecology & Evolution*, **35**, 56–67. <https://doi.org/10.1016/j.tree.2019.08.006>
- Issa N, Muller Y (2015) Atlas des oiseaux de France métropolitaine: Nidification et présence hivernale. DELACHAUX, Paris.
- IUCN (2001) IUCN Red List categories and criteria, version 3.1. IUCN.
- James G, Witten D, Hastie T, Tibshirani R (2013) *An Introduction to Statistical Learning: with Applications in R*. Springer, New York.
- Jarzyna MA, Jetz W (2016) Detecting the Multiple Facets of Biodiversity. *Trends in Ecology & Evolution*, **31**, 527–538. <https://doi.org/10.1016/j.tree.2016.04.002>
- Jetz W, McGeoch MA, Guralnick RP, Ferrier S, Beck J, Costello MJ, Fernandez M, Miguel Fernandez, Miguel Fernandez, Fernández M, Geller GN, Keil P, Merow C, Meyer C, Muller-Karger FE, Pereira HM, Regan EC, Regan E, Schmeller DS, Turak E (2019) Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology and Evolution*, **3**, 539–551. <https://doi.org/10.1038/s41559-019-0826-1>

- Jiguet F, Devictor V, Julliard R, Couvet D (2012) French citizens monitoring ordinary birds provide tools for conservation and ecological sciences. *Acta Oecologica*, **44**, 58–66. <https://doi.org/10.1016/j.actao.2011.05.003>
- Johnson CN, Balmford A, Brook BW, Buettel JC, Galetti M, Guangchun L, Wilmshurst JM (2017) Biodiversity losses and conservation responses in the Anthropocene. *Science*, **356**, 270–275. <https://doi.org/10.1126/science.aam9317>
- Johnson DS, Laake JL, Ver Hoef JM (2010) A Model-Based Approach for Making Ecological Inference from Distance Sampling Data. *Biometrics*, **66**, 310–318. <https://doi.org/10.1111/j.1541-0420.2009.01265.x>
- Johnston A, Hochachka WM, Strimas-Mackey ME, Gutierrez VR, Robinson OJ, Miller ET, Auer T, Kelling ST, Fink D (2019) Best practices for making reliable inferences from citizen science data: case study using eBird to estimate species distributions. *bioRxiv*, 574392. <https://doi.org/10.1101/574392>
- Johnston A, Matechou E, Dennis EB (2022) Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*, **14**, 103–116. <https://doi.org/10.1111/2041-210X.13834>
- Johnston A, Newson SE, Risely K, Musgrove AJ, Massimino D, Baillie SR, Pearce-Higgins JW (2014) Species traits explain variation in detectability of UK birds. *Bird Study*, **61**, 340–350. <https://doi.org/10.1080/00063657.2014.941787>
- Julliard R, Clavel J, Devictor V, Jiguet F, Couvet D (2006) Spatial segregation of specialists and generalists in bird communities. *Ecology Letters*, **9**, 1237–1244. <https://doi.org/10.1111/j.1461-0248.2006.00977.x>
- Kamp J, Oppel S, Heldbjerg H, Nyegaard T, Donald PF (2016) Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark. *Diversity and Distributions*, **22**, 1024–1035. <https://doi.org/10.1111/ddi.12463>
- Keller V, Herrando S, Voříšek P, Franch M, Kipson M, Milanese P, Martí D, Anton M, Klvaňová A, Kalyakin MV, Bauer H-G, Foppen RPB (2020) *European Breeding Bird Atlas 2: Distribution, Abundance and Change*. European Bird Census Council & Lynx Edicions, Barcelona.
- Kellner KF, Smith AD, Royle JA, Kéry M, Belant JL, Chandler RB (2023) The unmarked R package: Twelve years of advances in occurrence and abundance modelling in ecology. *Methods in Ecology and Evolution*, **14**, 1408–1415. <https://doi.org/10.1111/2041-210X.14123>
- Kellner KF, Swihart RK (2014) Accounting for Imperfect Detection in Ecology: A Quantitative Review. *PLOS ONE*, **9**, e111436. <https://doi.org/10.1371/journal.pone.0111436>
- Kéry M (2011) Towards the modelling of true species distributions. *Journal of Biogeography*, **38**, 617–618. <https://doi.org/10.1111/j.1365-2699.2011.02487.x>
- Kéry M, Gardner B, Monnerat C (2010) Predicting species distributions from checklist data using site-occupancy models. *Journal of Biogeography*, **37**, 1851–1862. <https://doi.org/10.1111/j.1365-2699.2010.02345.x>
- Kéry M, Royle JA (2015) Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models. Academic Press, Amsterdam ; Boston.

- Kéry M, Royle JA (2021) Chapter 12 - Conclusions. In: *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS* (eds Kéry M, Royle JA), pp. 717–723. Academic Press. <https://doi.org/10.1016/B978-0-12-809585-0.00012-0>
- Kéry M, Royle JA, Schmid H (2005) Modeling Avian Abundance from Replicated Counts Using Binomial Mixture Models. *Ecological Applications*, **15**, 1450–1461.
- Kéry M, Schmid H (2004) Monitoring programs need to take into account imperfect species detectability. *Basic and Applied Ecology*, **5**, 65–73. <https://doi.org/10.1078/1439-1791-00194>
- King R (2014) Statistical Ecology. *Annual Review of Statistics and Its Application*, **1**, 401–426. <https://doi.org/10.1146/annurev-statistics-022513-115633>
- Kullenberg C, Kasperowski D (2016) What Is Citizen Science? – A Scientometric Meta-Analysis. *PLOS ONE*, **11**, e0147152. <https://doi.org/10.1371/journal.pone.0147152>
- Lehikoinen A (2013) Climate change, phenology and species detectability in a monitoring scheme. *Population Ecology*, **55**, 315–323. <https://doi.org/10.1007/s10144-012-0359-9>
- Li M, Bolker B (2019) First release of phylogenetic comparative analysis in lme4-verse. <https://doi.org/10.5281/zenodo.2639887>
- Link WA, Schofield MR, Barker RJ, Sauer JR (2018) On the robustness of N-mixture models. *Ecology*, **99**, 1547–1551. <https://doi.org/10.1002/ecy.2362>
- Matutini F, Baudry J, Pain G, Sineau M, Pithon J (2021) How citizen science could improve species distribution models and their independent assessment. *Ecology and Evolution*, **11**, 3028–3039. <https://doi.org/10.1002/ece3.7210>
- Mayaud N, Balsac HH de, Jouard H (1936) *Inventaire des oiseaux de France*. Société d'études ornithologiques.
- McKinley DC, Miller-Rushing AJ, Ballard HL, Bonney R, Brown H, Cook-Patton SC, Evans DM, French RA, Parrish JK, Phillips TB, Ryan SF, Shanley LA, Shirk JL, Stepenuck KF, Weltzin JF, Wiggins A, Boyle OD, Briggs RD, Chapin SF, Hewitt DA, Preuss PW, Soukup MA (2017) Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation*, **208**, 15–28. <https://doi.org/10.1016/j.biocon.2016.05.015>
- McNew LB, Handel CM (2015) Evaluating species richness: Biased ecological inference results from spatial heterogeneity in detection probabilities. *Ecological Applications*, **25**, 1669–1680. <https://doi.org/10.1890/14-1248.1>
- Mesgaran MB, Cousens RD, Webber BL (2014) Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Diversity and Distributions*, **20**, 1147–1159. <https://doi.org/10.1111/ddi.12209>
- Michel NL, Smith AC, Clark RG, Morrissey CA, Hobson KA (2016) Differences in spatial synchrony and interspecific concordance inform guild-level population trends for aerial insectivorous birds. *Ecography*, **39**, 774–786. <https://doi.org/10.1111/ecog.01798>
- Mizel JD, Schmidt JH, Lindberg MS (2018) Accommodating temporary emigration in spatial distance sampling models. *Journal of Applied Ecology*, **55**, 1456–1464. <https://doi.org/10.1111/1365-2664.13053>

- Morin DJ, Yackulic CB, Diffendorfer JE, Lesmeister DB, Nielsen CK, Reid J, Schaub EM (2020) Is your ad hoc model selection strategy affecting your multimodel inference? *Ecosphere*, **11**, e02997. <https://doi.org/10.1002/ecs2.2997>
- Moussy C, Gwenaël Q, Gaudard C (2022) COMPTAGE DES OISEAUX D'EAU À LA MI-JANVIER EN FRANCE : Résultats 2022 du comptage Wetlands International. WETLANDS INTERNATIONAL.
- Navarro LM, Fernández N, Guerra CA, Rob Guralnick, Guralnick RP, Kissling WD, Londoño MC, Muller-Karger FE, Turak E, Balvanera P, Costello MJ, Delavaud A, Serafy GYE, Ferrier S, Geijzendorffer IR, Geller GN, Jetz W, Kim ES, Kim H, Martin CS, McGeoch MA, Mwampamba TH, Nel JL, Nicholson E, Pettorelli N, Schaeppman ME, Skidmore AK, Sousa-Pinto I, Pinto IS, Vergara SG, Vihervaara P, Xu H, Yahara T, Gill MJ, Pereira HM (2017) Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability*, **29**, 158–169. <https://doi.org/10.1016/j.cosust.2018.02.005>
- Neeson TM, Doran PJ, Ferris MC, Fitzpatrick KB, Herbert M, Khoury M, Moody AT, Ross J, Yacobson E, McIntyre PB (2018) Conserving rare species can have high opportunity costs for common species. *Global Change Biology*, **24**, 3862–3872. <https://doi.org/10.1111/gcb.14162>
- Newman G, Wiggins A, Crall A, Graham E, Newman S, Crowston K (2012) The future of citizen science: emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment*, **10**, 298–304. <https://doi.org/10.1890/110294>
- Nichols JD, Koneff MD, Heglund PJ, Knutson MG, Seamans ME, Lyons JE, Morton JM, Jones MT, Boomer GS, Williams BK (2011) Climate change, uncertainty, and natural resource management. *The Journal of Wildlife Management*, **75**, 6–18. <https://doi.org/10.1002/jwmg.33>
- Nichols JD, Thomas L, Conn PB (2009) Inferences About Landbird Abundance from Count Data: Recent Advances and Future Directions. In: *Modeling Demographic Processes In Marked Populations* Environmental and Ecological Statistics. (eds Thomson DL, Cooch EG, Conroy MJ), pp. 201–235. Springer US, Boston, MA. [https://doi.org/10.1007/978-0-387-78151-8\\_9](https://doi.org/10.1007/978-0-387-78151-8_9)
- Norvell RE, Howe FP, Parrish JR (2003) A Seven-Year Comparison of Relative-Abundance and Distance-Sampling Methods. *The Auk*, **120**, 1013–1028. <https://doi.org/10.1093/auk/120.4.1013>
- Paradis E, Baillie SR, Sutherland WJ, Gregory RD (1998) Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, **67**, 518–536. <https://doi.org/10.1046/j.1365-2656.1998.00215.x>
- Parmesan C (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology*, **13**, 1860–1872. <https://doi.org/10.1111/j.1365-2486.2007.01404.x>
- Payne EH, Gebregziabher M, Hardin JW, Ramakrishnan V, Egede LE (2018) An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data. *Communications in statistics: Simulation and computation*, **47**, 1722–1738. <https://doi.org/10.1080/03610918.2017.1323223>

- Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Carlo H. R. Heip, Carlo H. R. Heip, Carlo H. R. Heip, Heip CHR, Robert Hoft, Robert Hoft, Höft R, Hurtt GC, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M (2013) Essential biodiversity variables. *Science*, **339**, 277–278. <https://doi.org/10.1126/science.1229931>
- Pillar VD, Duarte L d. S (2010) A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters*, **13**, 587–596. <https://doi.org/10.1111/j.1461-0248.2010.01456.x>
- Pilotto F, Kühn I, Adrian R, Alber R, Alignier A, Andrews C, Bäck J, Barbaro L, Beaumont D, Beenaerts N, Benham S, Boukal DS, Bretagnolle V, Camatti E, Canullo R, Cardoso PG, Ens BJ, Everaert G, Evtimova V, Feuchtmayr H, García-González R, Gómez García D, Grandin U, Gutowski JM, Hadar L, Halada L, Halassy M, Hummel H, Huttunen K-L, Jaroszewicz B, Jensen TC, Kalivoda H, Schmidt IK, Kröncke I, Leinonen R, Martinho F, Meesenburg H, Meyer J, Minerbi S, Monteith D, Nikolov BP, Oro D, Ozoliņš D, Padedda BM, Pallett D, Pansera M, Pardal MÂ, Petriccione B, Pipan T, Pöyry J, Schäfer SM, Schaub M, Schneider SC, Skuja A, Soetaert K, Sprinĝe G, Stanchev R, Stockan JA, Stoll S, Sundqvist L, Thimonier A, Van Hoey G, Van Ryckegem G, Visser ME, Vorhauser S, Haase P (2020) Meta-analysis of multidecadal biodiversity trends in Europe. *Nature Communications*, **11**, 3486. <https://doi.org/10.1038/s41467-020-17171-y>
- Quaintenne G, Gaudard C, Béchet A, Benmergui M, Bouteaux J-J, Cadiou B, Camberlein P, Chapalain F, Croset F, Culioli J-M, Dalloyau S, Debout G, Dubois P, Dulac P, Flitti A, Gallien F, Gendre N, Girard O, Havet S, Vincent-Martin N (2020) Les oiseaux nicheurs rares et menacés en France en 2016 et 2017. Rare and endangered breeding bird survey in France in 2016-2017. , **Ornithos 27-2**, 73–111.
- Quinn JE, Brandle JR, Johnson RJ, Tyre AJ (2011) Application of detectability in the use of indicator species: A case study with birds. *Ecological Indicators*, **11**, 1413–1418. <https://doi.org/10.1016/j.ecolind.2011.03.003>
- Rabinowitz D (1981) Seven forms of rarity. *The biological aspects of rare plant conservation*, 205–217.
- Richter A, Nakamura G, Agra Iserhard C, da Silva Duarte L (2021) The hidden side of diversity: Effects of imperfect detection on multiple dimensions of biodiversity. *Ecology and Evolution*, **11**, 12508–12519. <https://doi.org/10.1002/ece3.7995>
- Ricotta C (2005) On parametric diversity indices in ecology: A historical note. *Community Ecology*, **6**, 241–242.
- Rigal S, Dakos V, Alonso H, Auniņš A, Benkő Z, Brotons L, Chodkiewicz T, Chylarecki P, de Carli E, del Moral JC, Domşa C, Escandell V, Fontaine B, Foppen R, Gregory R, Harris S, Herrando S, Husby M, Ieronymidou C, Jiguet F, Kennedy J, Klvaňová A, Kmecl P, Kuczyński L, Kurlavičius P, Kålås JA, Lehikoinen A, Lindström Å, Lorrillière R, Moshøj C, Nellis R, Noble D, ESKILDSEN DP, Paquet J-Y, Péliissié M, Pladevall C, Portolou D, Reif J, Schmid H, Seaman B, Szabo ZD, Szép T, Florenzano GT, Teufelbauer N, Trautmann S, van Turnhout C, Vermouzek Z, Vikstrøm T, Voříšek P, Weiserbs A, Devictor V (2023) Farmland practices are driving bird population decline

- across Europe. *Proceedings of the National Academy of Sciences*, **120**, e2216573120. <https://doi.org/10.1073/pnas.2216573120>
- Rigal S, Knape J (2024) Investigating the heterogeneity within Wild bird indices in Europe. *Biological Conservation*, **290**, 110452. <https://doi.org/10.1016/j.biocon.2024.110452>
- Robinson OJ, Socolar JB, Stuber EF, Auer T, Berryman AJ, Boersch-Supan PH, Brightsmith DJ, Burbidge AH, Butchart SHM, Davis CL, Dokter AM, Di Giacomo AS, Farnsworth A, Fink D, Hochachka WM, Howell PE, La Sorte FA, Lees AC, Marsden S, Martin R, Martin RO, Masello JF, Miller ET, Moodley Y, Musgrove A, Noble DG, Ojeda V, Quillfeldt P, Royle JA, Ruiz-Gutierrez V, Tella JL, Yorio P, Youngflesh C, Johnston A (2022) Extreme uncertainty and unquantifiable bias do not inform population sizes. *Proceedings of the National Academy of Sciences*, **119**, e2113862119. <https://doi.org/10.1073/pnas.2113862119>
- Roché J-E, Muller Y, Sibley J-P (2013) Une méthode simple pour estimer les populations d'oiseaux communs nicheurs en France. *Alauda*, **81**, 241–268.
- Santini L, Tobias JA, Callaghan C, Gallego-Zamorano J, Benítez-López A (2023) Global patterns and predictors of avian population density. *Global Ecology and Biogeography*, **32**, 1189–1204. <https://doi.org/10.1111/geb.13688>
- 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**, 1747–1765. <https://doi.org/10.1007/s10531-020-01948-0>
- Schaub M, Ullrich B (2021) A drop in immigration results in the extinction of a local woodchat shrike population. *Animal Conservation*, **24**, 335–345. <https://doi.org/10.1111/acv.12639>
- Schmidt JH, McIntyre CL, MacCluskie MC (2013) Accounting for incomplete detection: What are we estimating and how might it affect long-term passerine monitoring programs? *Biological Conservation*, **160**, 130–139. <https://doi.org/10.1016/j.biocon.2013.01.007>
- Schmidt JH, Rattenbury KL (2018) An open-population distance sampling framework for assessing population dynamics in group-dwelling species. *Methods in Ecology and Evolution*, **9**, 936–945. <https://doi.org/10.1111/2041-210X.12932>
- Skirvin AA (1981) EFFECT OF TIME OF DAY AND TIME OF SEASON ON THE NUMBER OF OBSERVATIONS AND DENSITY ESTIMATES OF BREEDING BIRDS. In: *Estimating Numbers of Terrestrial Birds* Studies in Avian Biology. (eds Ralph CJ, Scott JM), pp. 271–274. Cooper Ornithological Society.
- Sollmann R, Gardner B, Chandler RB, Royle JA, Sillett TS (2015) An open-population hierarchical distance sampling model. *Ecology*, **96**, 325–331. <https://doi.org/10.1890/14-1625.1>
- Sólymos P, Matsuoka SM, Stralberg D, Barker NKS, Bayne EM (2018) Phylogeny and species traits predict bird detectability. *Ecography*, **41**, 1595–1603. <https://doi.org/10.1111/ecog.03415>
- Suzuki-Ohno Y, Yokoyama J, Nakashizuka T, Kawata M (2017) Utilization of photographs taken by citizens for estimating bumblebee distributions. *Scientific Reports*, **7**, 11215. <https://doi.org/10.1038/s41598-017-10581-x>

- Thierion V, Vincent A, Valero S (2022) Theia OSO Land Cover Map 2020. <https://doi.org/10.5281/zenodo.6538861>
- Thorson JT, Barnes CL, Friedman ST, Morano JL, Siple MC (2023) Spatially varying coefficients can improve parsimony and descriptive power for species distribution models. *Ecography*, **2023**, e06510. <https://doi.org/10.1111/ecog.06510>
- Tingley MW, Beissinger SR (2013) Cryptic loss of montane avian richness and high community turnover over 100 years. *Ecology*, **94**, 598–609. <https://doi.org/10.1890/12-0928.1>
- Tredennick AT, Hooker G, Ellner SP, Adler PB (2021) A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, **102**, e03336. <https://doi.org/10.1002/ecy.3336>
- Tukey J (1977) *Exploratory Data Analysis*. Pearson, Reading, Mass.
- UICN France, MNHN, LPO, SEOF, ONCFS (2016) La Liste rouge des espèces menacées en France - Chapitre Oiseaux de France métropolitaine. Paris, France.
- Villéger S, Mason NWH, Mouillot D (2008) New Multidimensional Functional Diversity Indices for a Multifaceted Framework in Functional Ecology. *Ecology*, **89**, 2290–2301. <https://doi.org/10.1890/07-1206.1>
- Visser ME, Holleman LJM, Gienapp P (2006) Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia*, **147**, 164–172. <https://doi.org/10.1007/s00442-005-0299-6>
- Whelan CJ, Şekercioğlu ÇH, Wenny DG (2015) Why birds matter: from economic ornithology to ecosystem services. *Journal of Ornithology*, **156**, 227–238. <https://doi.org/10.1007/s10336-015-1229-y>
- Williams BK (2003) Policy, Research, and Adaptive Management in Avian Conservation. *The Auk*, **120**, 212–217. <https://doi.org/10.2307/4090161>
- Yeatman L, Ridet Y, d'ornithologie S française, France D de la protection de la nature (1976) *Atlas des oiseaux nicheurs de France de 1970 à 1975*. Société ornithologique de France.
- Yeatman-Berthelot D, Jarry G (1995) *Nouvel atlas des oiseaux nicheurs de France 1985-1989*. Societe Ornithologique De Fran, Paris.
- Zipkin EF, Inouye BD, Beissinger SR (2019) Innovations in data integration for modeling populations. *Ecology*, **100**, 1–3.
- Zipkin EF, Zylstra ER, Wright AD, Saunders SP, Finley AO, Dietze MC, Itter MS, Tingley MW (2021a) Addressing data integration challenges to link ecological processes across scales. *Frontiers in Ecology and the Environment*, **19**, 30–38. <https://doi.org/10.1002/fee.2290>
- Zipkin EF, Zylstra ER, Wright AD, Saunders SP, Finley AO, Dietze MC, Itter MS, Tingley MW (2021b) Addressing data integration challenges to link ecological processes across scales. *Frontiers in Ecology and the Environment*, **19**, 30–38. <https://doi.org/10.1002/fee.2290>



## IV Chapitre 2 : Développement de modèle intégré dans le but d'améliorer le suivi de l'avifaune nicheuse

---

### IV.A Contexte

Les données de sciences participatives constituent un afflux important d'informations permettant d'améliorer la surveillance des populations sauvages. Le volontariat de cette pratique de collecte occasionne un compromis entre (i) la collecte d'une information de qualité, *i.e.* donnée standardisée qui prospecte de façon homogène l'étendue spatiale et temporelle afin d'assurer la représentativité des habitats et de la fenêtre temporelle de la reproduction ; et (ii) la collecte d'une grande quantité d'informations, au détriment de l'assurance de la représentativité (Devictor, Whittaker & Beltrame, 2010).

L'estimation d'abondance non biaisée par le processus d'observation nécessite un ensemble d'informations permettant de dissocier les différentes composantes du processus d'observation (Nichols, Thomas & Conn, 2009; King, 2014). Les répliques temporels et distances d'observations collectés par le protocole EPOC-ODF permettent d'estimer, respectivement, la probabilité qu'un individu se rende disponible à l'observation ( $P_a$ ) et la probabilité que les individus soient détectés ( $P_d$ ), voir *Le vivant caché par l'observation*. Cependant, on constate que sur la période de collecte 2021-2023, le protocole EPOC a permis la collecte d'environ trois fois plus de points d'écoute que le protocole EPOC-ODF (Tableau 1) sur une plus grande couverture spatiale (Figure 26).

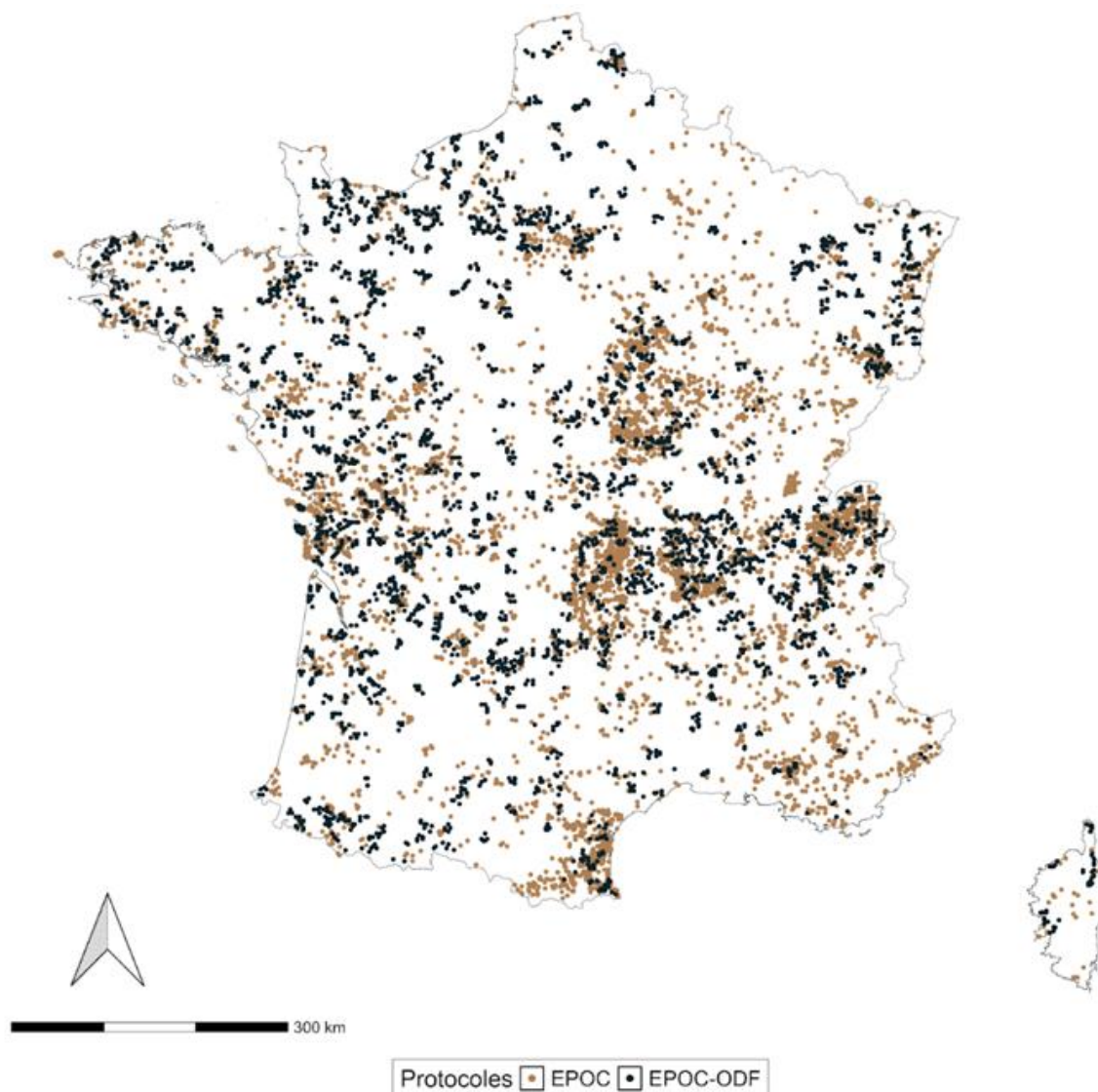


Figure 26 : Distribution spatiale des localités (listes et sites) des dispositifs EPOC et EPOC-ODF entre 2021 et 2023.

C'est dans ce contexte que pour le second chapitre, nous avons développé un modèle intégré de distance sampling (IDS), basé sur les travaux de Chandler, Royle & King, 2011 et Kéry et al., 2024, afin d'utiliser conjointement les protocoles EPOC et EPOC-ODF. Dans ce chapitre, nous avons réalisé une étude de simulation et une étude de cas sur trois régions françaises (Bourgogne-Franche-Comté, Normandie et Nouvelle-Aquitaine) illustrant un gradient en termes de proportions de données EPOC-ODF et EPOC collectées. L'étude de simulation a pour objectif de s'assurer que le modèle développé est en capacité d'estimer

correctement un ensemble de paramètres régissant le processus écologique et le processus d'observation des espèces selon différents gradients d'échantillonnage. De son côté, l'étude de cas a pour objectif de vérifier sur des données réelles, si le recours à l'intégration des jeux de données permet d'améliorer les estimations, *i.e.* réduire l'incertitude, selon différentes répartitions de données structurées/semi-structurées.

Les résultats de ce chapitre nous montrent que le recours à l'intégration des jeux de données permet de réduire l'incertitude des estimations. De plus, cette réduction d'incertitude n'est pas uniquement restreinte aux paramètres de distribution spatiale, mais elle a également des répercussions sur les paramètres liés au processus d'observation.

**Référence du contexte**

- Chandler RB, Royle JA, King DI. 2011. Inference about density and temporary emigration in unmarked populations. *Ecology* 92:1429–1435. DOI: 10.1890/10-2433.1.
- Devictor V, Whittaker RJ, Beltrame C. 2010. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions* 16:354–362. DOI: 10.1111/j.1472-4642.2009.00615.x.
- Kéry M, Royle JA, Hallman T, Robinson WD, Strebel N, Kellner KF. 2024. Integrated distance sampling models for simple point counts. *Ecology* n/a:e4292. DOI: 10.1002/ecy.4292.
- King R. 2014. Statistical Ecology. *Annual Review of Statistics and Its Application* 1:401–426. DOI: 10.1146/annurev-statistics-022513-115633.
- Nichols JD, Thomas L, Conn PB. 2009. Inferences About Landbird Abundance from Count Data: Recent Advances and Future Directions. In: Thomson DL, Cooch EG, Conroy MJ eds. *Modeling Demographic Processes In Marked Populations*. Environmental and Ecological Statistics. Boston, MA: Springer US, 201–235. DOI: 10.1007/978-0-387-78151-8\_9.

## IV.B Article

# **Improving national-scale Breeding Bird Surveys with Integrated Distance Sampling** *En révision dans Scientific Reports*

Jean Nabias<sup>1,2,\*</sup>, Romain Lorrillière<sup>2,3</sup>, Jérémy Dupuy<sup>1</sup>, Laurent Couzi<sup>1</sup>, Luc Barbaro<sup>2,4</sup>

1: LPO-BirdLife France, Fonderies Royales, Rochefort Cedex, France

2: CESCO, Muséum National d'Histoire Naturelle, CNRS, Sorbonne-University, Paris,  
France

3: Centre de Recherches sur la Biologie des Populations d'Oiseaux (CRBPO), MNHN-  
CNRS-OFB, Paris, France

4: Dynafor, INRA-INPT, University of Toulouse, Auzeville, France

Keywords : Bird monitoring ; Citizen science ; Distance sampling ; Data integration ;  
Hierarchical modelling ; Temporary emigration

## Abstract

Estimating bird population sizes over large spatial and temporal scales has been a long-time quest in ornithological science. To date, ornithologists mainly relied on the number of individuals detected in standard point counts, with derived estimates inferred from the proportionality assumption that actual population size represents a constant proportion of sampled individuals. However, providing realistic estimates of species abundance requires modelling the observation process with temporally replicated data, which is not always possible with the increasing use of ever-bigger datasets from citizen science programs. Data integration methods allow combining temporally replicated sampling at coarser spatial grains with data collected over larger spatial extents. Here, we used Integrated Distance Sampling (IDS) to combine national structured and semi-structured citizen-based bird surveys in France to estimate species abundances using observation distances and accounting for temporary emigration, *i.e.*, the probability of individuals being detectable during sampling visits. While our simulation study showed an overall increase in estimated model accuracy for both ecological and observation processes with available data, without significant bias, the case study suggests that these improvements depend on specific sampling scenarios. Integrated models represent a promising tool for ecological science, permitting joint use of poorly structured datasets with scale-restricted structured surveys.

## Introduction

Monitoring programs rely on the Essentials Biodiversity Variables (EBVs) framework, which includes indices such as species abundance, distributions and taxonomic or genetic diversities to assess biodiversity (Pereira et al., 2013; Proença et al., 2017). To estimate bird species abundance, ornithologists mainly rely upon standardised point-count methods

consisting of individual records of detected birds, either visual or auditory, over a given time period (Blondel et al., 1981). While these observed counts can be extrapolated to unsampled novel conditions through covariate relationships, using standard models such as GLMs (Generalised Linear Models), these models are based on the assumption that the number of detected individuals corresponds to a constant proportion of the sampled population across space and time (Yoccoz et al., 2001; Thompson, 2002). However, recent studies have shown that this assumption is not viable, due to variation in either species detectability or actual proportion of detected species (false negative ratio), both stemming from observation error per se (Kellner & Swihart, 2014) or because of recently changing conditions due to climate change (Lehikoinen, 2013).

Furthermore, due to logistical and financial constraints, data collection is inherently limited over standardisation and spatial scale trade-offs, even for rather well-monitored taxa such as breeding birds in temperate countries (Devictor et al., 2010). Given the nature and volume of data collected by standard protocols, bird ecologists must address that issue relying increasingly on more or less opportunistic or semi-structured Citizen Science (CS) programs (Castagneyrol et al., 2020). To overcome these concerns impeding further standardizations of population estimates (Nabias et al., 2024), the use of repeated visits and other collection methods, such as reports of detection distance or multiple observers, allow the combined modelling of the distinct ecological and observation processes (Guillera-Arroita, 2017), see Box 1.

While the ecological process corresponds to species response to environmental covariate variations through space and/or time, the observation process depicts data collection (Royle & Dorazio, 2008). Nichols et al. (2009) describe the observation process as the interaction of four components; (i) the probability that the individuals home ranges overlap the sampling units  $p_s$ ; (ii) given  $p_s$ , the probability that the individuals are present on the sampling units

during observers visits  $p_p$ ; (iii) the probability that individuals are available for detection (for instance, bird vocalisation during observers visits) denoted  $p_a$  and (iv) the probability of detection given individuals presence and detectability  $p_d$ . While  $p_s$  is assessed through sampling design and  $p_d$  can be inferred from specific data collection, such as detection distances;  $p_a$  and  $p_p$  need temporal replicates to be assessed and permitted to alleviate false negative errors (type II), considering a species absent at a site when undetected during list completion (Guillera-Arroita, 2017).

Ecological estimates, while explicitly accounting for these two distinct processes, require flexible statistical tools such as hierarchical models to model the ecological process of interest while accounting for the measurement error stemming from the observation process (Kéry & Royle, 2015). There is a large range of adequate statistical tools available depending on the ecological processes studied (King, 2014) from species presence/absence – (Occupancy models; MacKenzie et al., 2002), species abundance – (Hierarchical Distance Sampling, Sollmann et al., 2016 ; N-mixture models, Chandler et al., 2011) and demographic parameters estimation – Cormack-Jolly-Seber models, (Gimenez et al., 2007).

However, these types of data-hungry models require large sample sizes (Kéry & Royle, 2020) given the multiple estimated parameters, while also entailing the need for specific data collection (King, 2014).

In the last two decades, CS has seen exponential growth (Sullivan et al., 2009) thanks to the development of several online databases such as eBird ([www.ebird.org](http://www.ebird.org)), iNaturalist ([www.inaturalist.org](http://www.inaturalist.org)) and GBIF ([www.gbif.org](http://www.gbif.org)) aiming to handle observation data collected by volunteers (Bonney et al., 2009) over increasingly longer temporal and larger spatial scales (Hochachka et al., 2012). These databases rely mostly on opportunistic data, information gathered without sampling design or focused taxa (Soroye et al., 2018). While the use of metadata and ad hoc filters can increase the value of collected data (Johnston et al.,

2021), they tend to lack specificities of structured surveys, including intra- and inter-year repeated visits (Bayraktarov et al., 2019; Johnston et al., 2022).

Data integration, or the simultaneous joint analysis of an ecological process using multiple datasets (Schaub & Abadi, 2011), has triggered a growing interest in recent years (Fletcher et al., 2019; Zipkin et al., 2021). It is used, for instance, in the case of complex ecological inference requiring different data sources, such as Integrated Population Models (IPM) needing to count data as well as nest monitoring and/or banding to infer population spatiotemporal variations and population growth parameters (Besbeas et al., 2002; Schaub, 2020); or to combine data collected at different spatial and/or temporal resolutions (Keil et al., 2014).

Here, we focus on data collected for breeding bird atlases, depicting known distribution and population size estimates using data collected over a short timeframe (Donald & Fuller, 1998). In France, the previous atlas (Issa & Muller, 2015) was based on a semi-quantitative method (Roché et al., 2013). Based on extrapolating spatially the bird species densities known from a few local areas, this approach did not formerly account for the detection process. It therefore possibly resulted in biased estimations of French breeding bird populations when these are compared to estimates inferred from a structured CS scheme EPOC-ODF (Structured Estimation of Common Bird Population Size, see Nabias et al., 2024). While structured schemes result in intensive data collection to collect high-quality data, they inherently tend to be conducted over a rather limited spatial extent. Semi-structured schemes aim at overcoming this issue to gather interpretable data while still enlisting the largest possible number of observers and associated field data (Kelling et al., 2019). For our study, we used datasets from both the structured CS scheme EPOC-ODF and the semi-structured CS scheme EPOC (Estimation of Common Bird Population Size), where one scheme allows inference of the detection process through repeated visits (*i.e.* double-

sampling design; Mackenzie & Royle, 2005) while the other focus on the collection of environmental data without repeated visits.

Recent studies have shown the potential of model integration on ecological inferences combining data from multiple data sources for occupancy modelling (Lauret et al., 2021; von Hirschheydt et al., 2023) and species abundance estimates (Kéry et al., 2024). Here, we relied on a joint likelihood approach (Miller et al., 2019) expanded from the Integrated Distance Sampling (IDS) formulation from (Kéry et al., 2024). While these authors formulated an IDS model integrating data from unreplicated distance sampling data using point count and detection/non-detection data, here, we aimed at calibrating an IDS model accounting for temporary emigration, enabling distinct estimation of  $p_d$  and  $p_a p_p$  (Mizel et al., 2018), by using structured (temporal replicates over a few number of sites) and semi-structured datasets (no temporal replicates over a large number of sites). We modelled species abundance using two formulations: a Hierarchical Distance Sampling (HDS) and an IDS to test if data integration could lead to ecological estimates improvement.

The present study includes two simulation studies as well as an additional case study. The simulation studies aim at (i) assessing model identifiability, its capabilities to correctly estimate upon simulated species parameters of IDS formulation over simulated structured and semi-structured datasets with and without information sharing through effort covariates (Simulation 1A-B) ; and (ii) assessing improvement of ecological estimates, *i.e.* shrinkage of uncertainty intervals, over simulated gradients of species availability, detectability and sampling scenarios by comparing parameter posterior distribution from a HDS model (calibrated solely on structured data) and an IDS model (calibrated over structured and semi-structured datasets: Simulation 2). The case study completes simulation 2 by using real data collected over three French regions corresponding to distinct data collection scenarios, to

determine whether data integration can lead to ecological estimates improvement without leading to biases or changes in mean distribution.

## Material and Methods

### *Simulated and case study datasets*

Each dataset consisted of a structured dataset, with temporal replicates, and a semi-structured dataset, with single-visits. Observed distance data (either simulated or real) correspond to detection distances between observers and detected individuals, in metres, over point transect. While we randomly generated habitat and effort covariates and the ratio of structured and semi-structured data for the simulation study, we relied on EPOC-ODF (Structured Estimation of Common Bird Population Size) and EPOC (Estimation of Common Bird Population Size) citizen science schemes data. These two schemes consist of 5-minute point count completed checklists, where observers are tasked to point locations of detected individuals on the mobile app NaturaList. The distances of observation between observers and detected individuals are measured through GIS (Geographic Information System) using observers GPS location.

*Hierarchical Distance Sampling (HDS) [Box 1]*

Hierarchical Distance sampling (HDS) seeks to model spatial variation of abundance while taking account of the detection process (Kéry and Royle, 2015). As Conventional Distance sampling assumes perfect detection on the line (Buckland et al., 2015), HDS can relax this assumption by assessing the probability that the individual is present and available for detection during survey occasions (Chandler et al., 2011) through multiple visits at the same site. Considering a population following Poisson distribution with mean  $\lambda_i$ , at each site  $i = 1, 2, \dots, I$  we have the local population  $M_i$ :

$$M_i \sim \text{Poisson}(\lambda_i)$$

Given multiple visits  $j$  ( $j = 1, 2, \dots, J$ ), at site  $i$ , the number of individuals available for detection  $N_{i,j}$  follow binomial outcomes from the local population  $M_i$  with a probability of being exposed to sampling, *i.e.* available,  $\varphi_{i,j}$ :

$$N_{i,j} \sim \text{Binomial}(M_i, \varphi_{i,j})$$

For each site  $i$  and visit  $j$ , observers measure the distance of observation between themselves and observed individuals. A vector of cell probabilities  $\pi_{i,j}$  derived from a detection function  $f$  (Buckland et al., 2015), assigns probabilities to distinct distance bins. Observation  $y_{i,j}$  can then be described as a multinomial outcome given the number of individuals available for detection and its distance ( $x_{i,j}$ ):

$$y_{i,j} \sim \text{Multinomial}(N_{i,j}, \pi_{i,j}), \text{ with } \pi_{i,j} = f(x_{i,j}, \sigma)$$

In our study, we considered a half-normal model, with a parameter ( $\sigma$ ) for the detection function.

Buckland, S.T., Rexstad, E.A., Marques, T.A., Oedekoven, C.S., 2015. Distance Sampling: Methods and Applications, Methods in Statistical Ecology. Springer International Publishing. <https://doi.org/10.1007/978-3-319-19219-2>  
 Chandler, R.B., Royle, J.A., King, D.I., 2011. Inference about density and temporary emigration in unmarked populations. *Ecology* 92, 1429–1435. <https://doi.org/10.1890/10-2433.1>  
 Kéry, M., Royle, J.A., 2015. Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models, 1st edition. ed. Academic Press, Amsterdam ; Boston.

*Simulation study 1: Model identifiability*

To assess model identifiability, we simulated 3000 datasets using a set of parameters  $\varphi_{0,DSopen}$  (species availability in the case of structured datasets),  $\varphi_{0,DS}$  (species availability in the case of semi-structured datasets),  $\sigma_0$  (detectability),  $\beta_0$  (mean abundance) and coefficients ( $\gamma$ ) and ( $\alpha$ ), over covariables ( $X_j, V_j, U_{i,j}$ ) effects on availability and detectability respectively, collected at sites  $i$  during visits  $j$ .

Each dataset comprised a simulation of 200 sites with 9 temporal replicates (structured dataset) and 1000 sites with single-visits (semi-structured dataset) over one season (Fig 1). For the data simulation of 1A, we considered that covariates ( $X_j, V_j, U_{i,j}$ ) did not affect species availability and detectability by considering ( $\gamma$ ) and ( $\alpha$ ) equal to 0. We included them in simulation 1B (Eq. 1) to assess if the transfer of information through data collected by the two datasets could impact the model identifiability.

$$\left\{ \begin{array}{l} \log(\lambda_i) = \beta_0 \\ \text{logit}(\varphi_{j,DSopen}) = \varphi_{0,DSopen} + \gamma * X_j \\ \text{logit}(\varphi_{j,DS}) = \varphi_{0,DS} + \gamma * V_j \\ \log(\sigma_{i,j}) = \sigma_0 + \alpha * U_{i,j} \end{array} \right. \quad (\text{Eq 1})$$

We used an altered version of the function `simHDSopen` from *AHMbook* (Kéry et al., 2023) to simulate the datasets. All models were calibrated using JAGS 4.3.1 (Plummer, 2003) through *jagsUI* (Kellner & Meredith, 2021) R package, while MCMC samples were retrieved using *mcmcoutput* (Juat et al., 2022). See appendix S1 for MCMC parameters and priors used for simulation and case study.

*Simulation study 2: Integrated datasets and sample sizes*

We assessed inference improvement through data integration by comparing estimated parameters obtained from a HDS model, calibrated using only structured datasets, with those estimated using the IDS formulation of simulation 1B. We used the same set of parameters

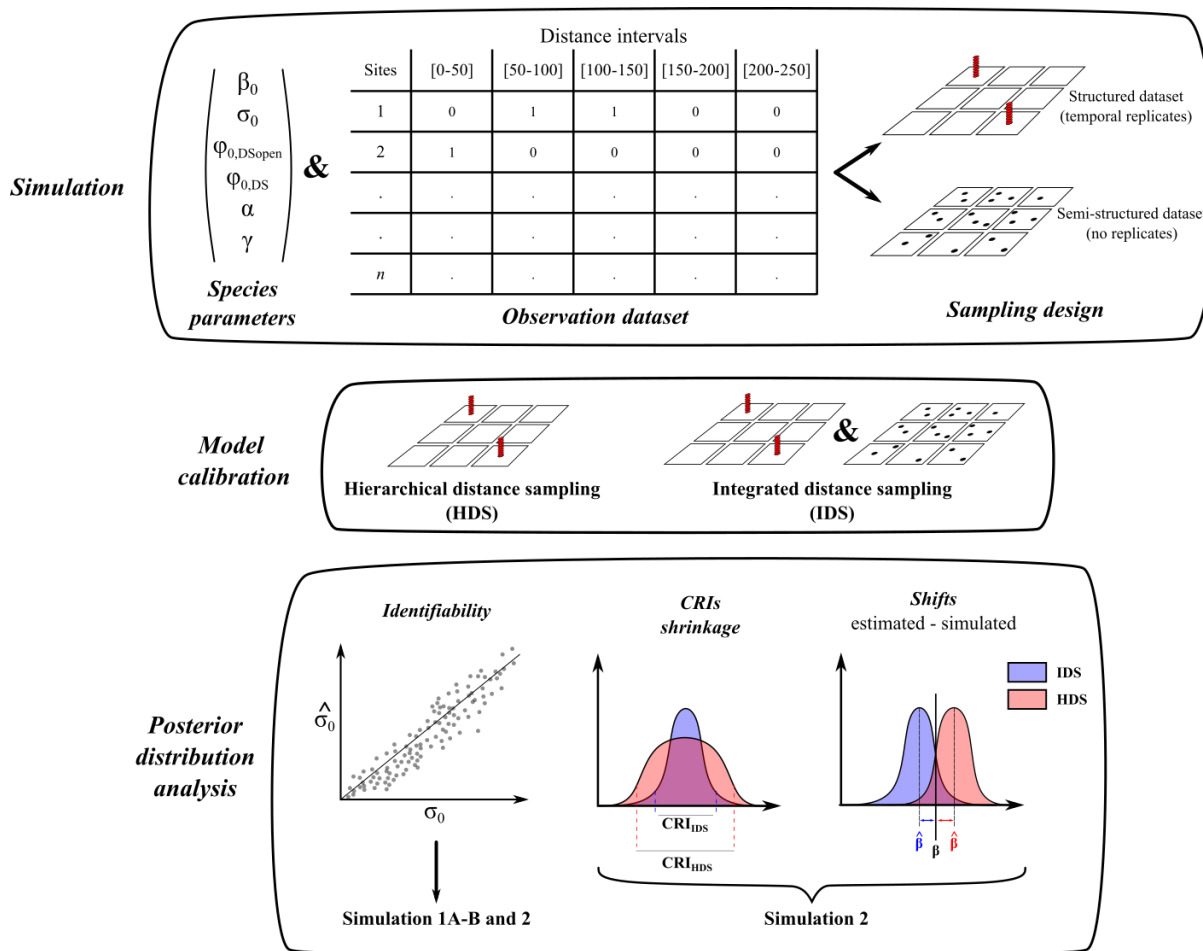
from simulation 1 while varying the number of sites of the structured dataset from 50 to 300 and the structured/semi-structured ratio surveys from 1 to 6. The choice for delimiting the number of structured sites and ratio is based on our case study, see figure 2 and appendix S2. Inference improvement was associated with a shrinkage of uncertainty (*i.e.* reduction of the span from the 95% posterior distribution of estimated parameters – Credible intervals CRIs). Here is the pseudo R-syntax of the model:

$$\log(\text{CRIs}) \sim \text{Model} + \text{Parameter} + \varphi_0 + \sigma_0 + \text{Nb of structured sites} + \text{Model}$$

*\* added ratio of semi structured data*

We calibrated a log-normal linear model for assessing CRIs shrinkage using model formulation used (*i.e.*, HDS or IDS), parameters estimated ( $\beta_0, \varphi_0, \text{DSopen}, \sigma_0, \gamma, \alpha$ ). As we expect that the model formulation could benefit from the number of input data, we also included an interaction between the model formulation and the simulated sampling design, *i.e.* the ratio of semi-structured sites.

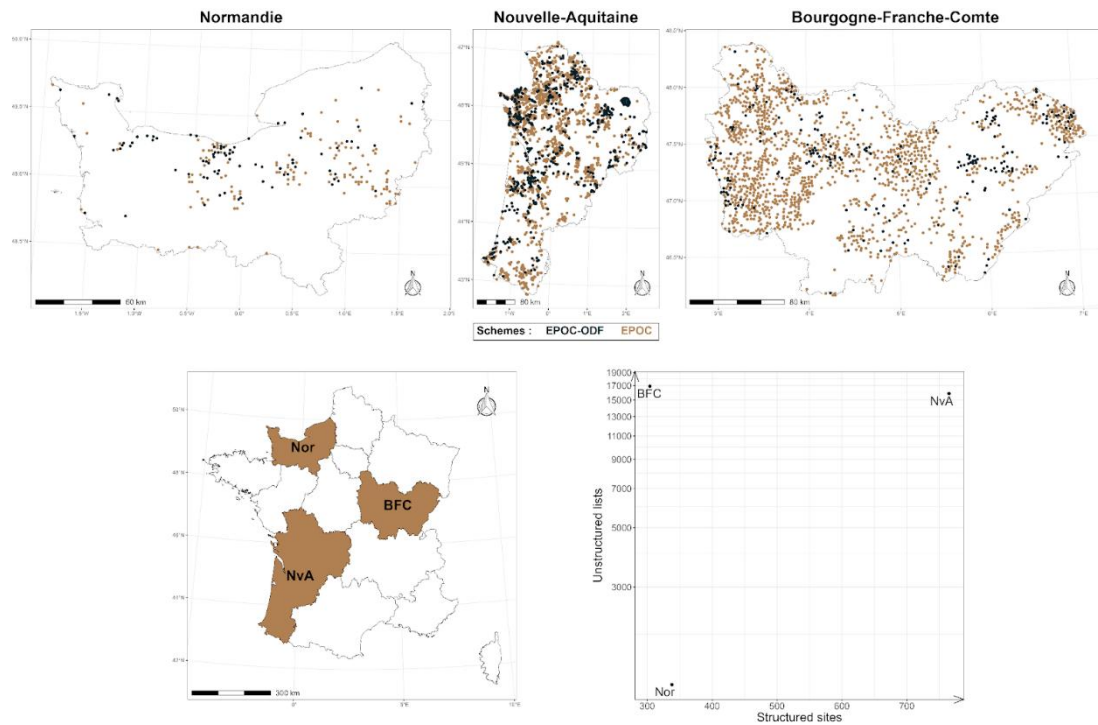
To assess robustness, we used a bootstrap incrementation analysis on the signal across various scenarios. Using all simulations resulting in converging models for both IDS and HDS formulation, we systematically varied the amount of information used to calibrate the log-linear model, by incrementing the number of simulation outputs in steps of 100. For each subset, we conducted a resample of size 300 to compute confidence intervals. See appendix S4 for more details and the robustness analysis.



**Fig 1:** Schematic representation of the simulation study design. We simulated 3000 datasets (structured and semi-structured) using the same set of simulated parameters across the two simulation studies. We simulated distinct detection probabilities  $\varphi_{0,DSopen}$  and  $\varphi_{0,DS}$  for each sampling design aiming to mimic a protocol effect. For the simulation study 1A, we didn't include covariable effect over availability ( $\gamma$ ) and detectability ( $\alpha$ ).

### Case study

We used data from 31 bird species collected during their breeding season over 2021-2023 across three French regions (Bourgogne-Franche-Comté, Nouvelle-Aquitaine and Normandie). These regional datasets differ in terms of data quantity, allowing to test multiple habitat gradients scattered in France while depicting a diverse distribution of structured and semi-structured data collections (Fig 2.).

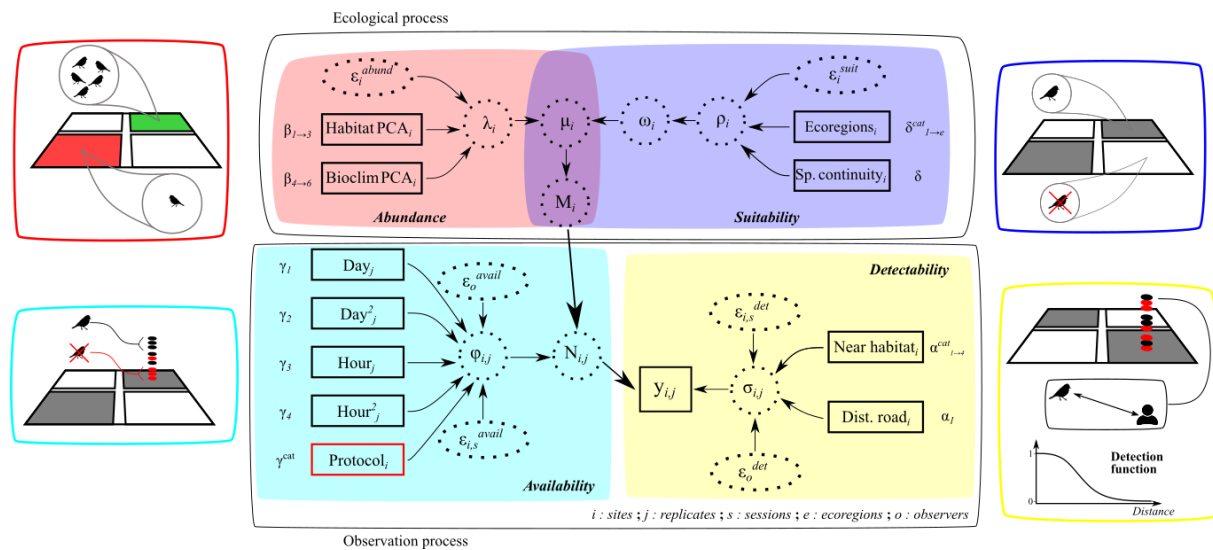


**Fig 2:** Spatial distribution of structured sites (EPOC-ODF) and semi-structured lists (EPOC) over selected French regions (Nor : Normandie ; NvA : Nouvelle-Aquitaine ; BFC : Bourgogne-Franche-Comté).

For the EPOC scheme, there are no spatial constraints nor obligation to carry repeated visits, whereas, for EPOC-ODF, survey locations are randomly selected from a systematic grid and have to be visited three times during the breeding season, each session consisting of three successive 5-minute point counts. For the semi-structured dataset (EPOC), we applied a spatial filter to select EPOC lists collected at least two kilometres away from sites with temporal replicates (EPOC-ODF) and other EPOC lists, see Appendix S2. For each species, we calibrated a HDS model, using only data collected by the EPOC-ODF schemes and an IDS model using data collected by both schemes.

Bird species selection was based upon targeted species from the two schemes (Nabias et al., 2024) and had a sufficient number of observations, at least detected once over 20 distinct EPOC-ODF sites, in each region. We applied a temporal filter that considered both observed bird activities during the breeding season and expert knowledge to define the breeding phenology of each targeted species and exclude possible early or late migrants. For

each targeted species, we applied a right-side truncation of 5% over the observation distance to remove extreme distance values for model robustness (Buckland et al., 2001).



**Fig 3:** Directed acyclic graph (DAG) representation of the model and its parameter. Observed and latent variables are represented using solid squares and dotted circles respectively. Arrows depicted links between parameters and covariates. Estimated coefficients are depicted on the side of covariates. We include the protocol origin as a categorical covariate, represented by a red box, solely for the IDS model. Each subprocess is represented by distinctive colours and pictograms, from left-to-right and up-to-down we depicted processes (i) describing variation of species abundance across space in relation to habitat covariates; (ii) representing sites' probability of being considered unsuitable for modelled specie; (iii) assessing species probability of being exposed to sampling occasions and (iv) depicting species probability of being detected given its observation distance.

We modelled the population of a site  $M_i$  using a Zero-inflated Poisson of parameter  $\mu_i$  (Fig 3):

$$M_i \sim \text{Poisson}(\mu_i), \text{ with } \mu_i = \lambda_i * (1 - \omega_i)$$

Species abundance parameter ( $\lambda_i$ ) was modelled using reduced habitat (Thierion et al., 2022) and bioclimatic (Fick & Hijmans, 2017) covariates obtained through PCA (Nabias et al., 2024).

$$\log(\lambda_i) = \beta_0 + \sum_{a=1}^3 \beta_a * \text{Habitat PCAs}_i + \sum_{a=4}^6 \beta_a * \text{Bioclimatic PCAs}_i + \varepsilon_i^{\text{abund}}$$

Zero-inflated Poisson component ( $\omega_i$ ) corresponds to site suitability depicted by a Bernoulli process with the probability ( $\rho_i$ ) of a being considered unsuitable. We modelled  $\rho_i$  in regards to site ecoregions  $e$ , as a categorical variable (IGN, 2011) and its spatial continuity

(Guetté et al., 2021). We also included a site random effect for abundance ( $\varepsilon_i^{abund}$ ) and site suitability ( $\varepsilon_i^{suit}$ ).

$$\omega_i \sim \text{Bernoulli}(\rho_i)$$

$$\text{logit}(\rho_i) = \rho_0 + \sum_{a=1}^e \delta_a^{cat} * \text{Ecoregion}_i + \delta * \text{Spatial continuity}_i + \varepsilon_i^{suit}$$

We modelled the probability of an individual being available for detection ( $\varphi_{i,j}$ ) using covariates such as hour from sunrise and Julian date with quadratic effect to represent birds phenology across the breeding season. For the IDS model, we included a categorical covariate regarding to the scheme associated with the observation ( $\gamma^{cat}$ ).

$$\begin{aligned} \text{logit}(\varphi_{i,j}) = & \varphi_0 + \gamma_1 * \text{Day}_j + \gamma_2 * \text{Day}_j^2 + \gamma_3 * \text{Hr. sun}_j + \gamma_4 * \text{Hr. sun}_j^2 + \varepsilon_{i,j}^{avail} \\ & + \varepsilon_0^{avail} \end{aligned}$$

For the species detectability, we used a half-normal detection function with parameter ( $\sigma_{i,j}$ ), where we modelled observers detection probabilities in regards to observed distances using categorical variables describing the habitat over four categories (Agricultural, Forest, Open and Urban; Nabias et al., 2024) as well as the distance between their positions and the nearest road (Cote et al., 2021). We considered two distinct intercepts for each dataset.

$$\log(\sigma_{i,j}) = \sigma_0 + \alpha * \text{Dist. Road}_i + \sum_{a=2}^4 \alpha_a^{cat} * \text{Near habitat}_i + \varepsilon_{i,j}^{det} + \varepsilon_0^{det}$$

For species availability and detectability; we accounted for the study design of the structured dataset by implementing random effects over each session ( $\varepsilon_{i,j}^{avail}$  and  $\varepsilon_{i,j}^{det}$ ) while also adding observers random effect, as one observer can partake in both CS schemes ( $\varepsilon_0^{avail}$  and  $\varepsilon_0^{det}$ ).

After retrieval of HDS and IDS models output, we calibrated a log-normal and a linear model for assessing CRIs shrinkage and shift in mean estimated parameters respectively. We calibrated linear mixed effect models using lme4 (Bates et al., 2015) to

include nested random effects over species and studied regions, while having estimated parameters and an interaction between model formulation and studied regions as fixed variables. Here is the pseudo R-syntax of the model:

$$\log(CRIs) \sim Model * Region + Parameters + (1|Species\Region)$$

We used *emmeans* (Lenth et al., 2024) to estimate marginal means for the log-linear model and pairwise post hoc multiple comparisons of estimated marginal means from the linear model. For the case study analysis, we removed ( $\gamma^{cat}$ ) and ( $\delta^{cat}$ ) parameters from comparison as for  $\gamma^{cat}$  isn't estimated for the HDS model and  $\delta^{cat}$  parameters varied across studied regions.

Out of the 31 studied species, we excluded four species: Ring-necked Pheasant (*Phasianus colchicus*), Eurasian Jay (*Garrulus glandarius*), Eurasian Blackcap (*Sylvia atricapilla*) and Eurasian Collared-Dove (*Streptopelia decaocto*); due to model non-convergence, *i.e.* one of monitored parameters with  $\hat{R} \leq 1.1$ , of either HDS or IDS model in at least one of the three regions.

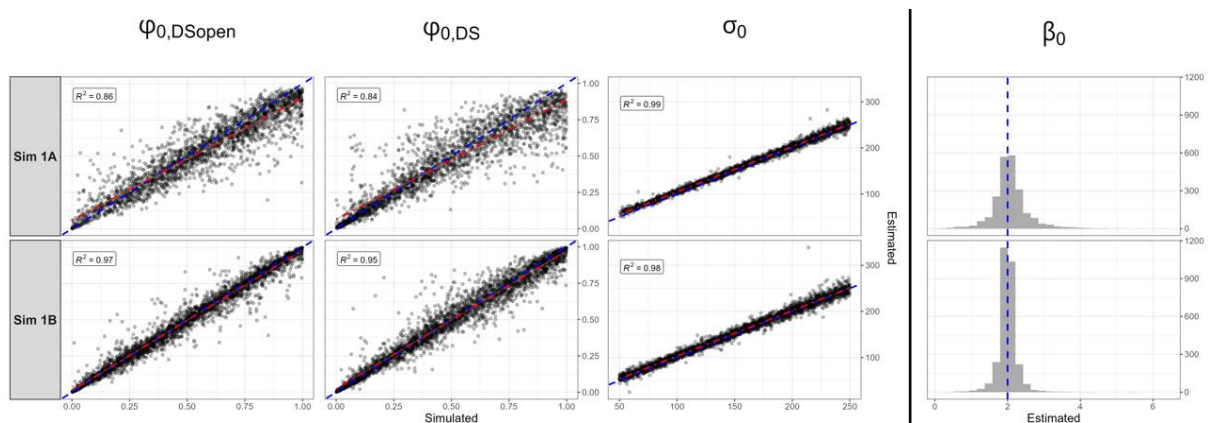
## Results

### *Simulation study*

#### *Simulation study 1*

For the first simulation study, out of 3000 simulated datasets, we had 2266 converging models (75.5%) for simulation 1A and 2865 (95.5%) for simulation 1B. When we include covariate coefficients and allow information sharing between the two datasets (0.86 and 0.84 for simulation 1A, 0.97 and 0.95 for simulation 1B; Fig 4), there is an increase in  $R^2$  values for  $\varphi_{0,DSopen}$  and  $\varphi_{0,DS}$ . There was no change in  $R^2$  value for  $\sigma$  between 1A and 1B, while  $\beta_0$  estimated parameters were more centred on the simulated values in simulation 1B

(Fig 4). IDS model calibrated for simulation 1B also showed signs of identifiability for coefficients  $\alpha$  and  $\gamma$  (Fig S3.1).



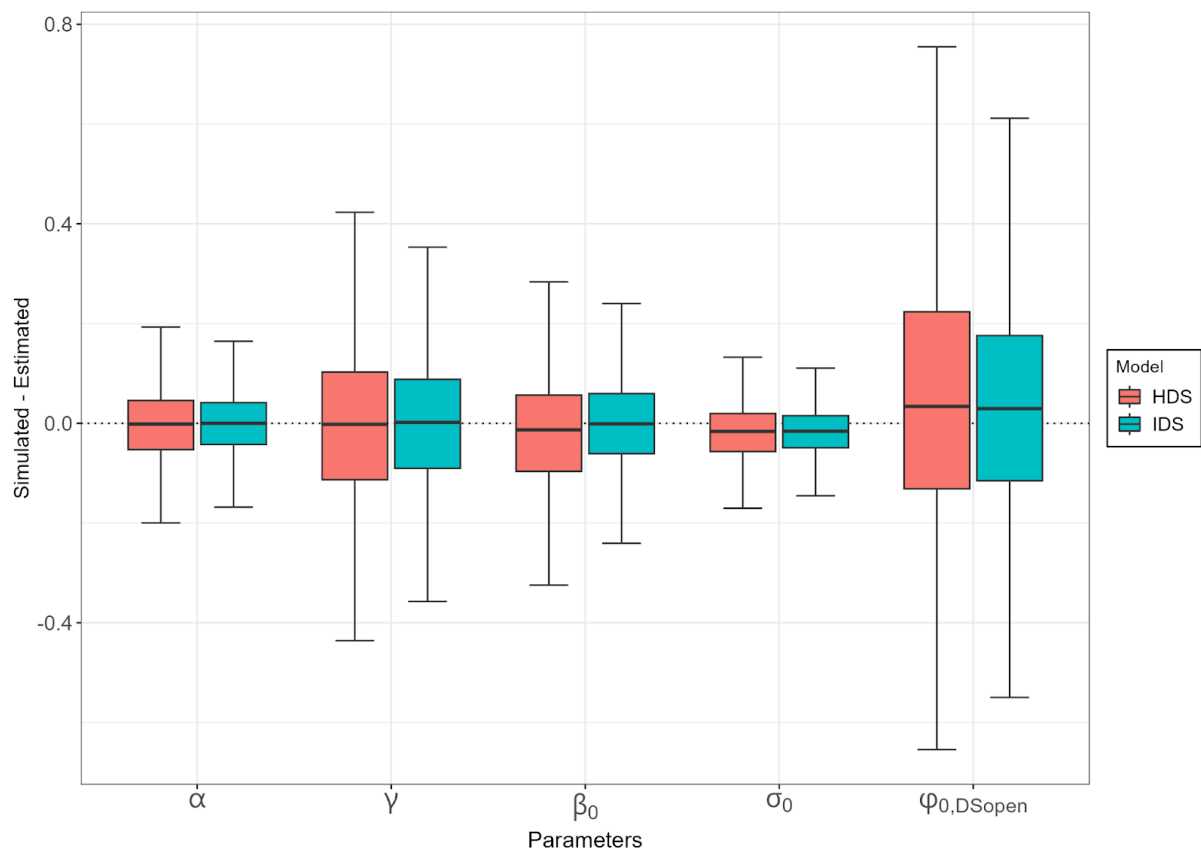
**Fig 4: Identifiability plot of converged model for the simulation study 1 for  $0$ ,  $\varphi_{0,DSopen}$ ,  $\varphi_{0,DS}$ ,  $\sigma_0$ ,  $\beta_0$  and their associated linear regression (dotted red line)  $R^2$  values. Results for the model without information sharing between the two datasets through  $\alpha$  and  $\gamma$  are shown in the first row (*i.e.* Sim 1A), while results for the model with shared information are shown in the second row (Sim 1B). Accurate parameter identification is represented by a dotted blue line. The  $\beta_0$  parameter is depicted as a histogram of estimated values, as we didn't vary it across simulations.**

### Simulation study 2

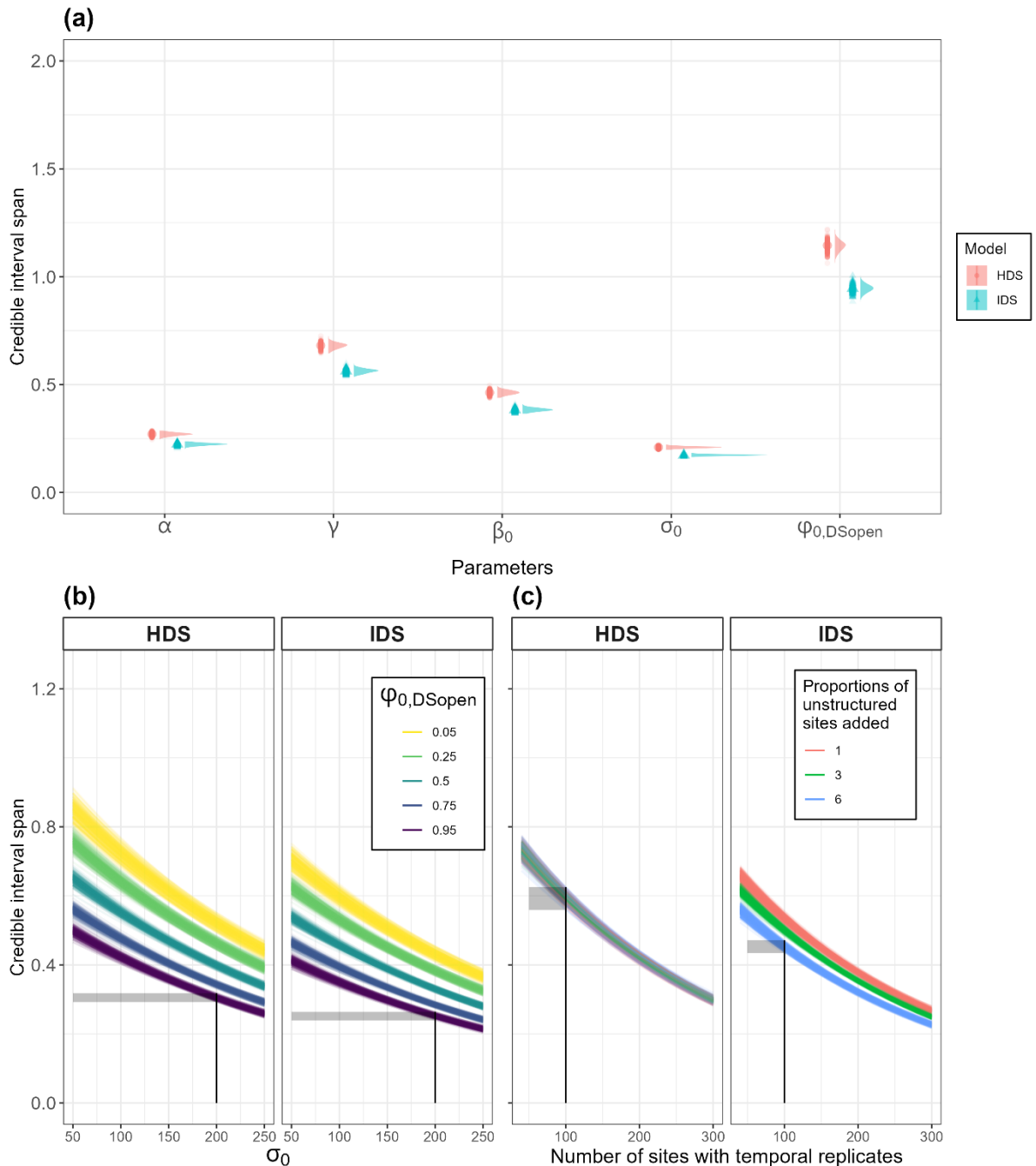
Out of 3000 simulated datasets, we had 2884 (96.1%) converging models using HDS formulation and 2877 (95.9%) converging models using IDS formulation. We see a slight increase of the  $R^2$  value for identifiability of  $\varphi_{0,DSopen}$  for IDS model and no changes for  $0$  and  $0$  (See appendix S3.2). IDS model shows a sign of identifiability for ( $\varphi_{0,DS}$ ;  $R^2 = 0.93$ ). There were no shifts between mean estimated parameters and their simulated counterpart for  $\beta_0$ ,  $\varphi_{0,DSopen}$ ,  $\sigma_0$ ,  $\alpha$  and  $\gamma$  (Fig 5).

The log-linear model was calibrated using output summaries from 2787 simulated datasets resulting in converging models for both HDS and IDS formulation. We obtained a considerable reduction of CRIs span across all estimated parameters for IDS models (Fig 6.a). Overall, IDS and HDS models produced narrower CRIs for available and more easily detectable species, however, the IDS model produced narrower CRIs, for equivalent species availability-detectability profiles simulated (Fig 6.b) than HDS. The number of structured

sites, *i.e.* including temporal replicates, was correlated with a reduction of CRI span for both models (Fig 6.c), although IDS model CRI reduction was also correlated with an increasing proportion of semi-structured sites added to the calibration dataset (Fig 6.c). While the increasing proportion of semi-structured sites added to the calibration dataset had no substantial effect on the HDS model, we found an important correlation to a CRIs reduction for the IDS model (Fig 6.c; see appendix S4 for robustness analysis of the signal across incrementation).



**Fig 5:** Boxplot of differences (simulated - estimated) across mean estimated parameters for HDS and IDS models. Each parameter is represented in their respective scale (log or logit). Accurate estimation, *i.e.* no shift, is depicted by the dotted line.



**Fig 6:** Marginal plot the simulation study 2 representing credible intervals (CRIs) span obtained from HDS and IDS models over 2787 simulations where both models converged. (a) Average CRIs and their associated bootstrapped confidence intervals (CIs), depicted by vertical density plots, over simulated parameters. (b) CRIs average responses and CIs over simulated species availability ( $\varphi_{0,DSopen}$ ) and detectability ( $\sigma_0$ ) continuums on natural scales. Species availability averaged across multiple classes (0.05,0.25,0.5,0.75 and 0.95 detection probability) are depicted by colour-graded lines. (c) CRIs average responses and CIs over multiple data collection cases. Proportions of semi-structured data (without temporal replicates) averaged over three classes (1,3 and 6) are depicted by colour-graded lines.

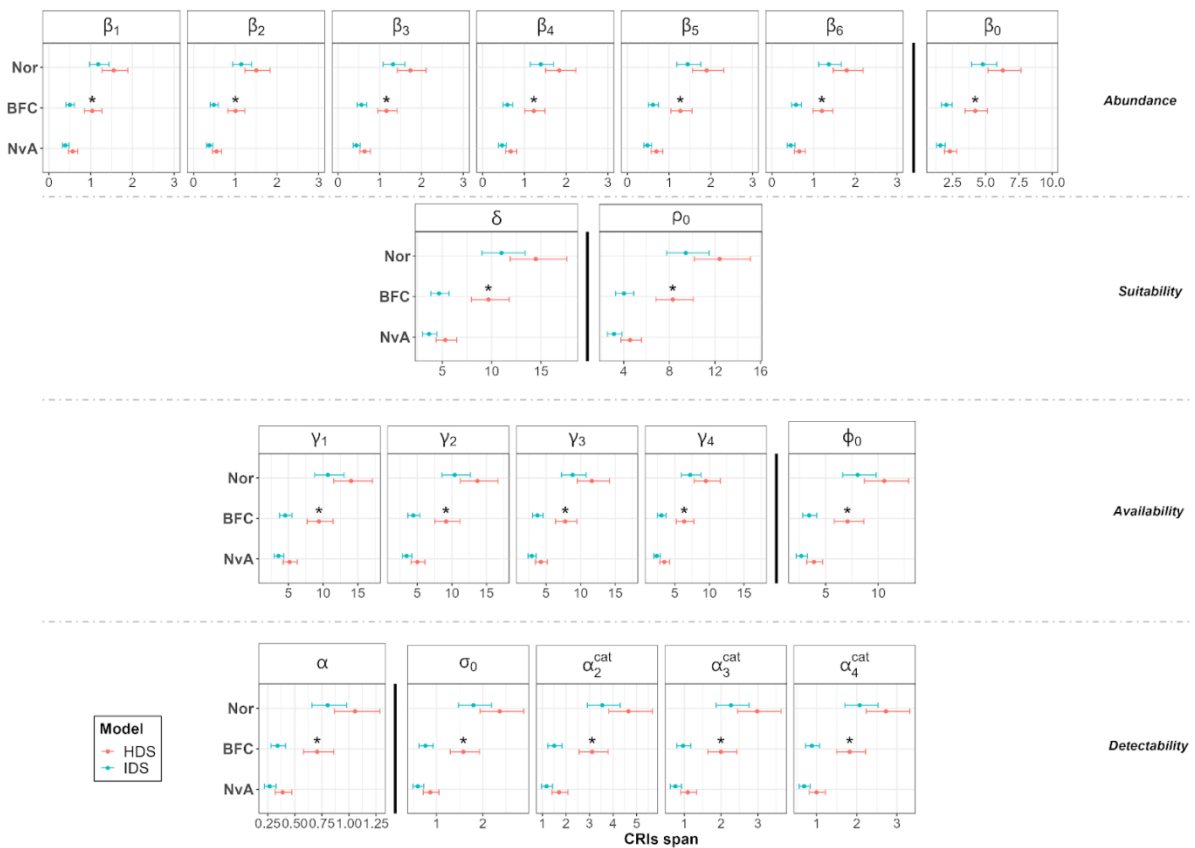
Each dot (a) and line (b-c) corresponds to a model response from a bootstrapped resample of the considered subset (number of used output data = 900, see appendix S4 for robustness analysis of the response signal), allowing visualisation of the response signal CIs.

For comparison between HDS and IDS estimates accuracy, we plotted CRIs responses, grey ribbons, for the case of simulated species with high mean detectability ( $\sigma_0 = 200$ m) and high probability of being available

$(\varphi_{0,DSopen} = 0.95; \mathbf{B})$  surveyed over 100 sites with temporal replicates and six times the number of semi-structured sites ( $\mathbf{c}$ ), depicted with vertical lines. The Lower and upper bounds of the rectangles correspond to minimal and maximal estimated means CRIs values.

### Case study

Marginal effect plots from the log-linear model (Fig 7) showed that Credible Intervals (CRIs) were slightly wider for Normandie (**Nor**), the region with less structured sites and semi-structured sites than Nouvelle-Aquitaine (**NvA**) and Bourgogne-Franche-Comté (**BFC**), regions with a larger number of semi-structured sites (Fig 7). While there were no significant differences (no overlap of marginal response confidence intervals) between HDS and IDS CRIs for **NvA** and **Nor**, CRIs from IDS calibration were significantly narrower than HDS ones for all estimated parameters in **BFC**. Pairwise comparison of marginal means showed no signs of significant ( $p$ -values  $> 0.5$ ) differences between HDS and IDS mean estimated parameters across all monitored parameters and studied regions. Squared-GVIFs (Generalised Variance-Inflation Factor; Fox & Monette, 1992), measured using *car* R package (Fox et al., 2023), are less than 4, showing no signs of multicollinearity for the terms used in each model.



**Fig 7:** Marginal effect plots of CRIs spans of estimated parameters and their associated CIs, in regards to the model (IDS in blue and HDS in red) and French regions (Nor: Normandie; NvA: Nouvelle-Aquitaine and BFC: Bourgogne-Franche-Comté). Results for parameters associated with site abundance, site suitability, species availability and species detectability are represented from top to bottom respectively. In each row, coefficients are depicted to the left of the vertical black line, while intercepts are represented on the right, for species detectability we depicted the intercept ( $\sigma_0$ ) as well as its associated categorical effects ( $\alpha_{2-4}$ ). Significant gaps of averaged mean estimated parameters between models are highlighted with an asterisk.

## Discussion

The present work brings new evidence that Integrated Distance Sampling (IDS) models can accurately identify parameters of a complex ecological process and expand their application accounting for temporary emigration. Moreover, it also shows that data integration improves ecological inference, through the shrinkage of credible intervals (CRI) span, for all parameters of the studied ecological process, across multiple sampling design scenarios and species availability-detectability continuums. The results from the case study further strengthen the simulation study, by showing that this shrinkage of CRI span without

significant variations of estimated mean parameters depends on the ratio of structured and semi-structured data used for each case study.

In recent years, there has been an increase in the interest for integrated models (Fletcher et al., 2019; Zipkin et al., 2021) due to their efficiency at reducing potential biases inherent to a single dataset (Zipkin et al., 2019) and allowing reliance on automated and non-invasive data collection methods (Doser et al., 2021; Pavanato Julião, 2021). It is worth noting that data integration through joint likelihood (Dietze, 2017) still has potential drawbacks when temporal and/or spatial mismatches, corresponding to discontinuity between dataset timeframes and spatial heterogeneity, are unaccounted for. Such mismatches could lead to biased inferences where the sampled timeframes and/or regions don't correctly represent the ecological process of interest (Powney et al., 2014; Isaac et al., 2020).

Here, we simulated an unbiased collection of environmental covariates across national-scale habitat gradients for the structured and semi-structured datasets, which can be misrepresentative of citizen science spatial sampling bias (Johnston et al., 2020). For the case study, we accommodated these mismatches by using data collected over the same timeframe, 2021-2023, and applying a spatial filter over the semi-structured dataset based upon the 2x2 km systematic grid resolution of the structured scheme. This resulted in an important decrease in available data from the semi-structured dataset (see Appendix S2) that could be resolved through random effect (Zipkin et al., 2021). In our case, we could consider distinct spatial subsets of the semi-structured dataset and implement them into a random effect structure encompassing the modelled sub-processes of the integrated model.

While hierarchical models offer a viable option to disentangle variations due to the observation process from the variations originating from the ecological process of interest (Guillera-Arroita, 2017), the trade-off between data specificities and data quantity can limit their applications. Data integration corresponds to a valuable option to increase the number of

available data to help calibrate such models. In this study, we integrate two datasets with similar sampling guidelines akin to a double-sampling design (Mackenzie & Royle, 2005). Data integration needs to account for integrated sampling schemes dissimilarities, e.g., variable species availability in regards to list duration between standardised schemes and non-standardised schemes with varying durations (Kéry et al., 2024). Options to calibrate integrated hierarchical models exist in a frequentist framework (Kéry et al., 2024) allowing fast computation. However, given the types of available data and ecological processes of interest, data integration is prone to rely on Bayesian frameworks. Bayesian computation is based on Markov Chain Monte Carlo (MCMC) techniques which are computationally high-demanding (Dorazio, 2016). Novel approaches exist such as Bayesian emulation or Integrated Nested Laplace Approximation (INLA – Rue et al., 2017; Fer et al., 2018) allowing efficient computation and facilitating implementation of spatial components (Blangiardo et al., 2013).

Integrated models could represent a great tool for macro-ecology related studies, spanning across large spatial scales (Zipkin et al., 2021) or requiring multiple institutions coordinating data collection (Navarro et al., 2017; Silva del Pozo et al., 2023). It could be used, for instance, in the study of bird populations across Europe from the Pan-European Common Bird Monitoring (PECBMS), which gathers data from point count, line transect, or territory mapping schemes across 28 countries and varying numbers of fieldworkers (Brlík et al., 2021), while taking account of country discrepancies in sampling design, sampling effort or varying starting period that could alter estimation of long-term trend (Duchenne et al., 2022). The joint analysis of multiple data sources could represent a substantial increase in data availability for the study of cryptic species (Martin et al., 2023; Twining et al., 2024), the use of data collected upon schemes lacking design-based methodology (Farr et al., 2024) and improve assessment of migratory patterns over large spatial scales (Meehan et al., 2022).

It represents an influx of data for the estimation of an ecological process of interest (Zipkin et al., 2021), that could potentially alleviate the sampling effort of robust design. However, we recommend carefully assessing, before their implementation, if 'lessen' robust sampling designs aiming at data integration are still capable of estimating targeted parameters of the whole ecological process or only partially, using power analysis (Guillera-Arroita & Lahoz-Monfort, 2012) and assessment of the overall integrated model identifiability via simulations (Ogle & Barber, 2020).

Our results highlight the benefits of relying on statistical frameworks such as Integrated Models capable of improving estimates accuracy through expansion of usable data collected from structured and semi-structured surveys. While our simulation results showed a constant reduction of estimates uncertainty, results from field surveys in three distinct French regions, depicting distinct ratios in quantity of structured and semi-structured data, showed that this improvement is case-dependant and significantly reduced estimates uncertainty with a low quantity of structured data and high quantity of semi-structured data. While we advocate for thorough planning before sampling, this suggests that Integrated Models could represent a conceivable alternative in case of insufficient collection from structured surveys and could also greatly benefit from data collected by citizen science schemes.

### Acknowledgement

EPOC and EPOC-ODF schemes are supervised by LPO (Ligue pour la Protection des Oiseaux), the representative of Birdlife in France. We thank all participants and local coordinators from regional instances for contributing to data collection through faune-france.org, a collaborative online database. We would like to express our thanks to Benoit Fontaine for his comments on prior versions of the manuscript. IDS and HDS models were calibrated on the SACADO MeSU platform at Sorbonne Université in Paris – France.

### Author contributions

All authors contributed to the current work. Study conceptualization was led by JN,RL,LB. Data collection management was performed by JD and LC. Analysis was carried out by JN and RL. Project administration was supervised by LB and LC. Writing was led by JN. Writing and reviews were conducted by LB, RL.

### Competing Interests

The authors declare no competing interests.

### Code and data availability

Scripts, BUGS model files and data for simulation and case studies replications are available online: <https://doi.org/10.5281/zenodo.11452854>

### Fundings

Funding for this work was provided through OFB (Office français de la Biodiversité), LPO and ANRT (Association Nationale Recherche Technologie; CIFRE grant, number: 2021/0305). The French Ministry of the Environment and OFB support the LPO through multi-year objectives agreements, in particular, to consolidate several EBV relating to birds, based on citizen science monitoring.

## IV.C Supplementary information

*S1: MCMC parameters and priors*

### ***Simulation studies (1A, 1B and 2)***

#### MCMC parameters

Iteration: 30.000

Burn-in: 10.000

Thin: 40

Chains: 3

#### Range for species parameter generation

$$\varphi_{0,DSopen} \sim U(0,1)$$

$$\varphi_{0,DS} \sim U(0,1)$$

$$\gamma \sim U(-2,2)$$

$$\alpha \sim U(-2,2)$$

#### Specific for the simulation study 2:

Number of structured sites (with temporal replicates)  $\sim U(50,250)$ , rounded to the first decimal.

Ratio of semi-structured lists and structured sites  $\sim U(1,6)$

#### MCMC Priors

$$\bar{\beta}_0 \sim U(0,10) \text{ and } \beta_0 = \log(\bar{\beta}_0)$$

$$\bar{\sigma}_0 \sim U(0,500) \text{ and } \sigma_0 = \log(\bar{\sigma}_0)$$

$$\bar{\varphi}_{0,DSopen} \sim U(0,1) \text{ and } \varphi_{0,DSopen} = \text{logit}(\bar{\varphi}_{0,DSopen})$$

$$\bar{\varphi}_{0,DS} \sim U(0,1) \text{ and } \varphi_{0,DS} = \text{logit}(\bar{\varphi}_{0,DS})$$

$$\alpha \sim N(0,10)$$

$$\gamma \sim N(0,10)$$

## ***Case study***

### MCMC parameters

Iteration: 350.000

Burn-in: 175.000

Thin: 35

Chains: 4

### Priors

#### Abundance

$$\beta_0 \sim N(0,10)$$

$$\tau_{\beta_0} \sim \text{Gamma}(1,1)$$

$$\varepsilon_i^{\text{abund}} \sim N(0, \tau_{\beta_0}), \text{ with } i \text{ corresponding to sites}$$

$$\beta_{1 \rightarrow 6} \sim N(0,10)$$

#### Suitability

$$\bar{\rho}_0 \sim U(0,1) \text{ and } \rho_0 = \text{logit}(\bar{\rho}_0)$$

$$\tau_{\rho_0} \sim \text{Gamma}(1,1)$$

$$\varepsilon_i^{\text{suit}} \sim N(0, \tau_{\rho_0})$$

$$\delta \sim N(0,10)$$

$$\delta_{1 \rightarrow e}^{\text{cat}} \sim N(0,10), \text{ with } e \text{ corresponding to the number of ecoregions}$$

#### Availability

$$\bar{\varphi}_0 \sim U(0,1) \text{ and } \varphi_0 = \text{logit}(\bar{\varphi}_0)$$

$$\gamma_{1 \rightarrow 4} \sim N(0,10)$$

$$\gamma^{\text{cat}} \sim N(0,10)$$

$$\varepsilon_o^{\text{avail}} \sim N(0, \tau_1^{\text{avail}}), \text{ with } o \text{ corresponding to observers}$$

$$\varepsilon_i^{avail} \sim N(0, \tau_2^{avail})$$

$$\varepsilon_{i,j}^{avail} \sim N(0, \tau_3^{avail}), \text{ with } j \text{ corresponding to visits occasions}$$

$$\tau_1^{avail}, \tau_2^{avail} \text{ and } \tau_3^{avail} \sim \text{Gamma}(1,1)$$

### Detectability

$$\bar{\sigma}_0 \sim U(0.1, 500) \text{ and } \sigma_0 = \log(\bar{\sigma}_0)$$

$$\alpha \sim N(0, 10)$$

$$\alpha_{1 \rightarrow 4}^{cat} \sim N(0, 10)$$

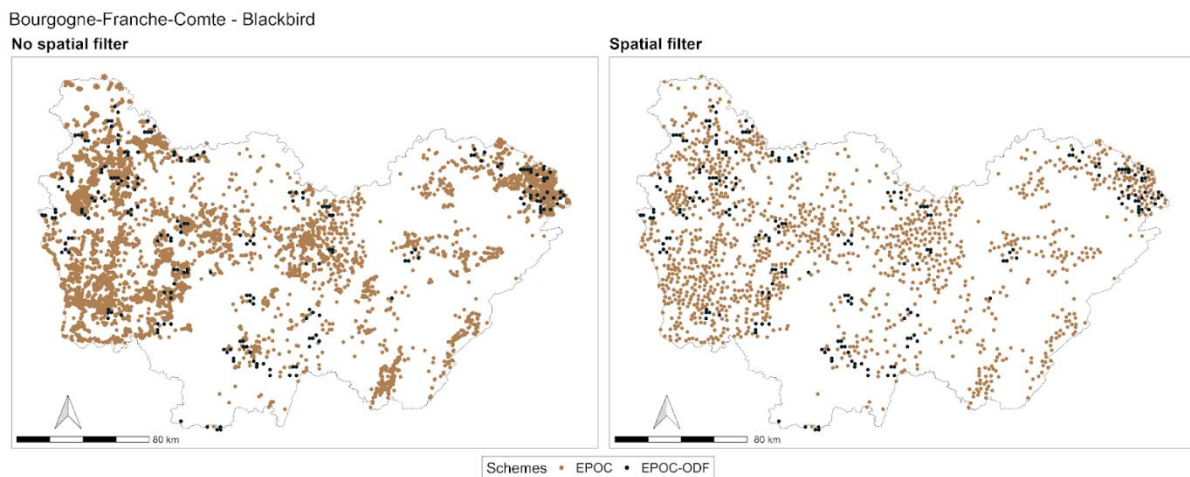
$$\varepsilon_0^{det} \sim N(0, \tau_1^{det})$$

$$\varepsilon_i^{det} \sim N(0, \tau_2^{detl})$$

$$\varepsilon_{i,j}^{det} \sim N(0, \tau_3^{det})$$

$$\tau_1^{det}, \tau_2^{det} \text{ and } \tau_3^{det} \sim \text{Gamma}(1,1)$$

S2: Semi-structured dataset spatial filter

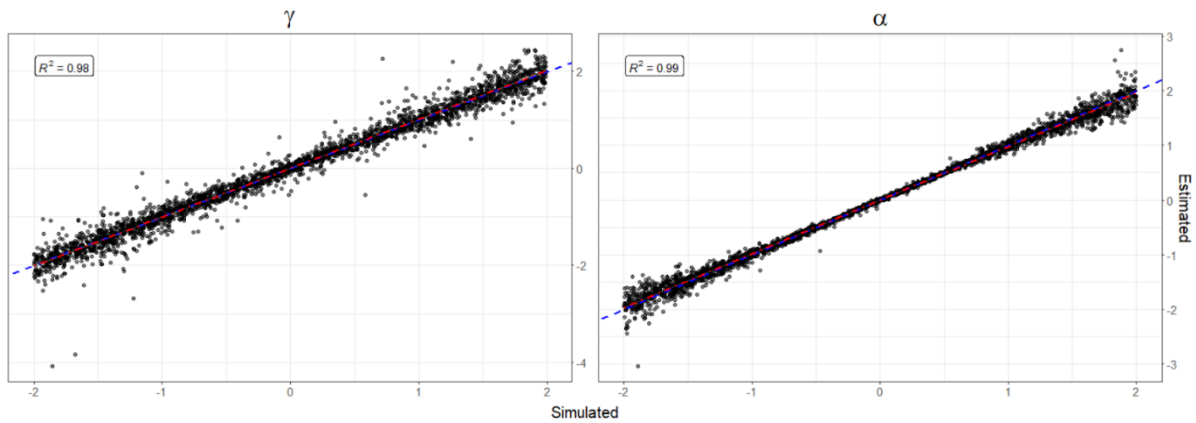


**Fig. S2.1:** Example of the spatial filter in Bourgogne-Franche-Comte region for the Blackbird (*Turdus merula*).

Observation data collected by birders across the breeding season spanning from March 1st to June 31th. For mountainous regions, a slight shift later in the season is considered (late March to late July). We first apply a filter over collected lists to target species phenology (available in reproducibility shared data). Spatial filters are specific of the semi-structured dataset (EPOC). We first select all EPOC lists over 2 kilometres distances of structured dataset (EPOC-ODF) sites. From this first selection, we subsampled EPOC lists spaced by at least 2 kilometres, the resolution of EPOC-ODF sampling design (Appendix S1, Nabias et al., 2024). Taking the example of the blackbird, prior to spatial filtering, we had 10,670 EPOC lists and 211 EPOC-ODF sites and ended up with 1,385 EPOC lists and 211 EPOC-ODF sites, a reduction of 87% of EPOC usable data.

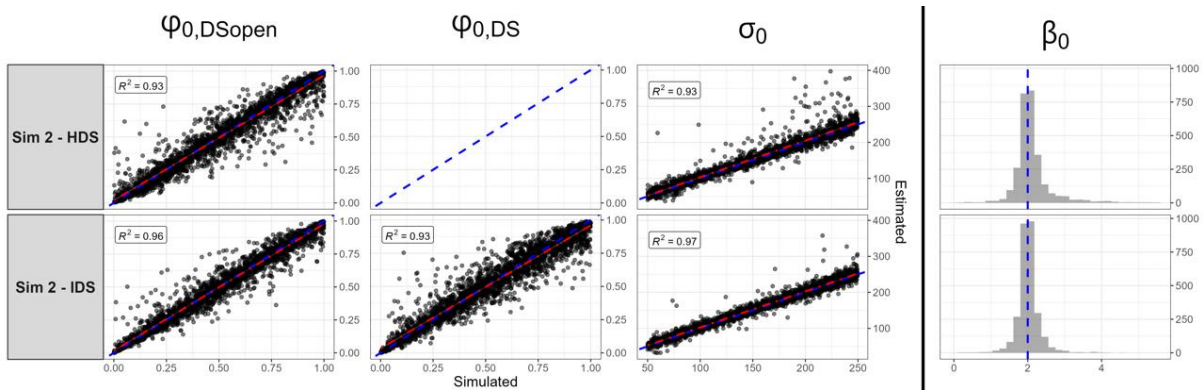
S3: Simulations identifiability: Complementary figures

**Simulation 1B: Coefficients identifiability**

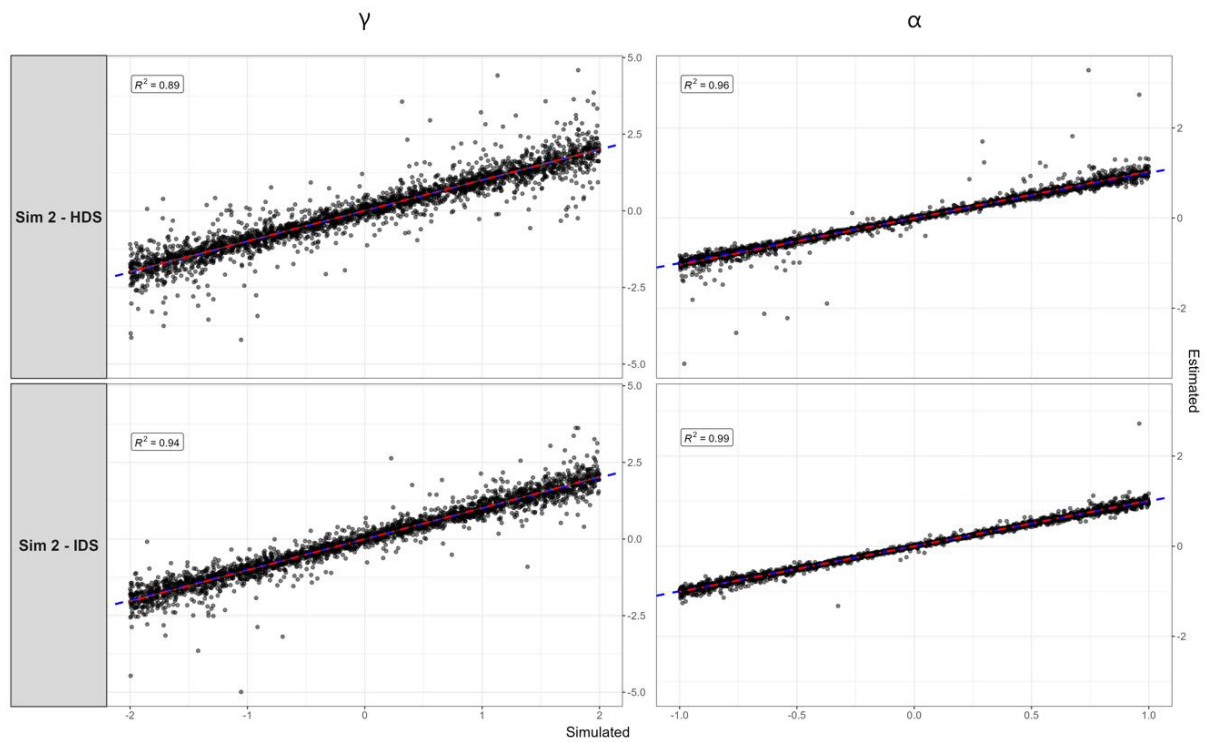


**Fig S3.1:** Simulation 1B identifiability plot for availability ( $\gamma$ ) and detectability ( $\alpha$ ) coefficients and their associated  $R^2$  values.

**Simulation 2: Identifiability**



**Fig S3.2:** Identifiability plot for the simulation study 2 for  $\varphi_{0,DSopen}$ ,  $\varphi_{0,DS}$ ,  $\sigma_0$ ,  $\beta_0$  and their associated linear regression (dotted red line)  $R^2$  values. Results from the model using only structured data (HDS) are shown in the first row, while results for the model using both structured and semi-structured data (IDS) are shown in the second row. Accurate parameter identification is represented by a dotted blue line. The  $\beta_0$  parameter is depicted as a histogram of estimated values, as we didn't vary it across simulation.



**Fig S3.3:** Simulation 2 identifiability plot for availability ( $\gamma$ ) and detectability ( $\alpha$ ) coefficients and their associated  $R^2$  values.

*S4: Simulations study 2: Robustness analysis*

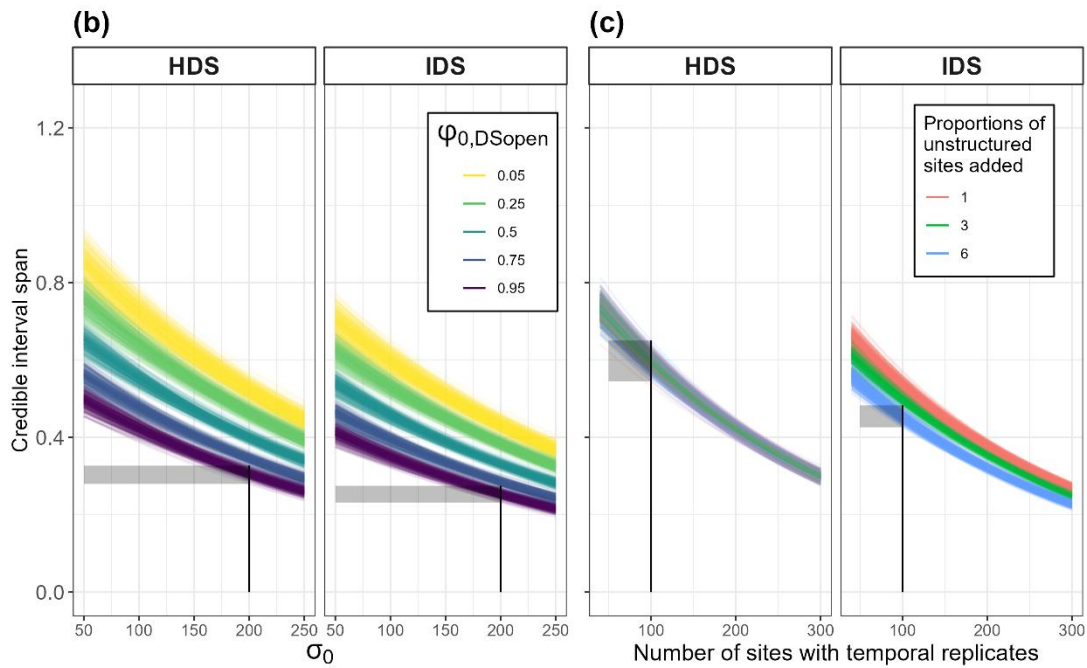
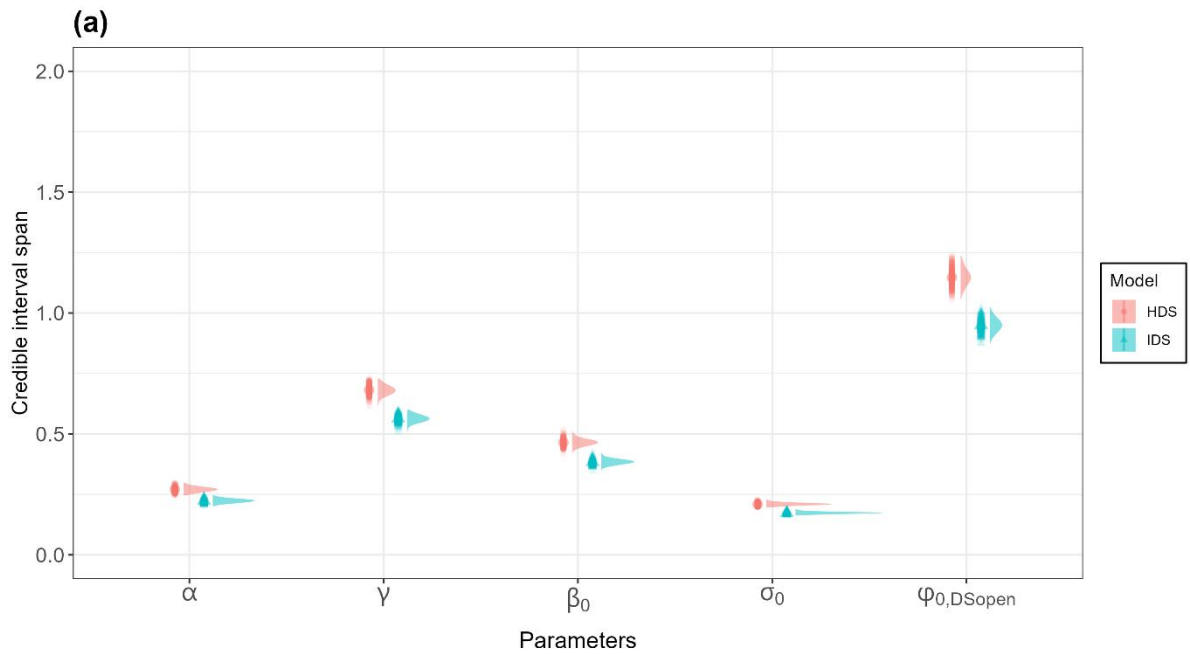
The robustness analysis was based upon converging models ( $n = 2,787$ ), for both HDS and IDS formulation. From this original pool of simulated parameters – estimated parameters posteriors distributions associations, we subsampled the quantity of data used for the log-linear model. Signal perseverance was assessed through incremental increases in the quantity of data used, using steps of 100 (total of 28 increments step, the final 28th step using all available data). We performed 300 resamples at each subset. For the two following figures, we plotted marginal plot, prediction of the Credible Intervals (CRIs) span response across specific parameters depicted in the main text, see section “Material and Methods - Simulation study 2: Integrated datasets and sample sizes”, while over were averaged. We used 2787 simulation outputs with converged IDS and HDS models.

These plots highlight signal perseverance of the CRI span diminution despite changes in the number of data included in our increment bootstrap analysis. We limited the number of available data to 500 (1/5th of total data) for the first plot and 1500 (3/5th of total data) for the second plot. As a reminder, figure 6, in the main text, used 900 output simulations for each resamples.

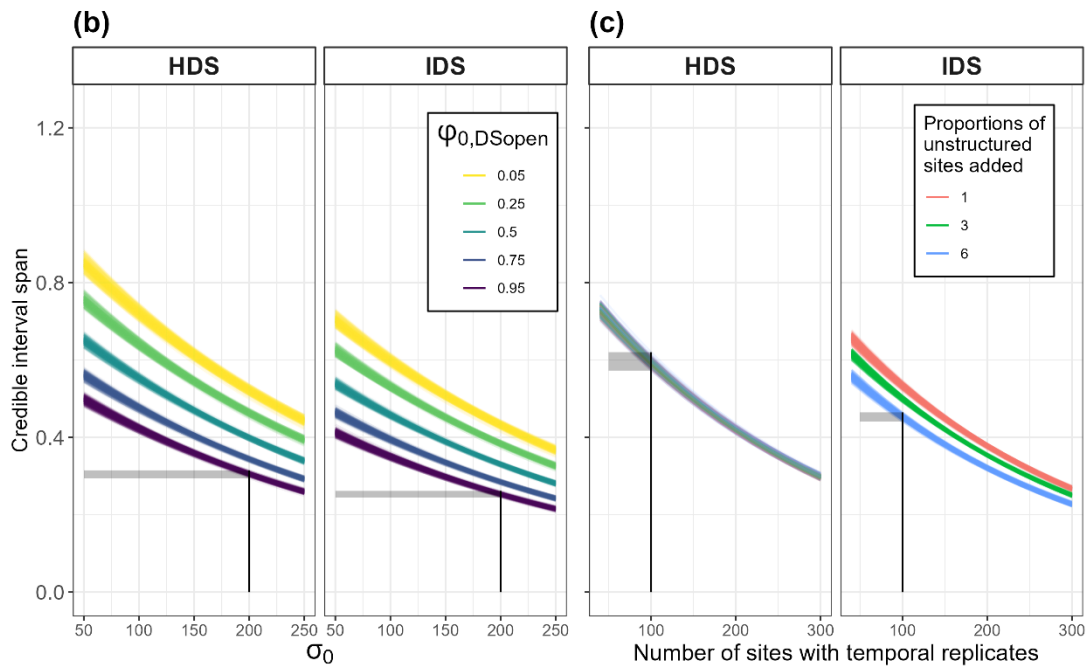
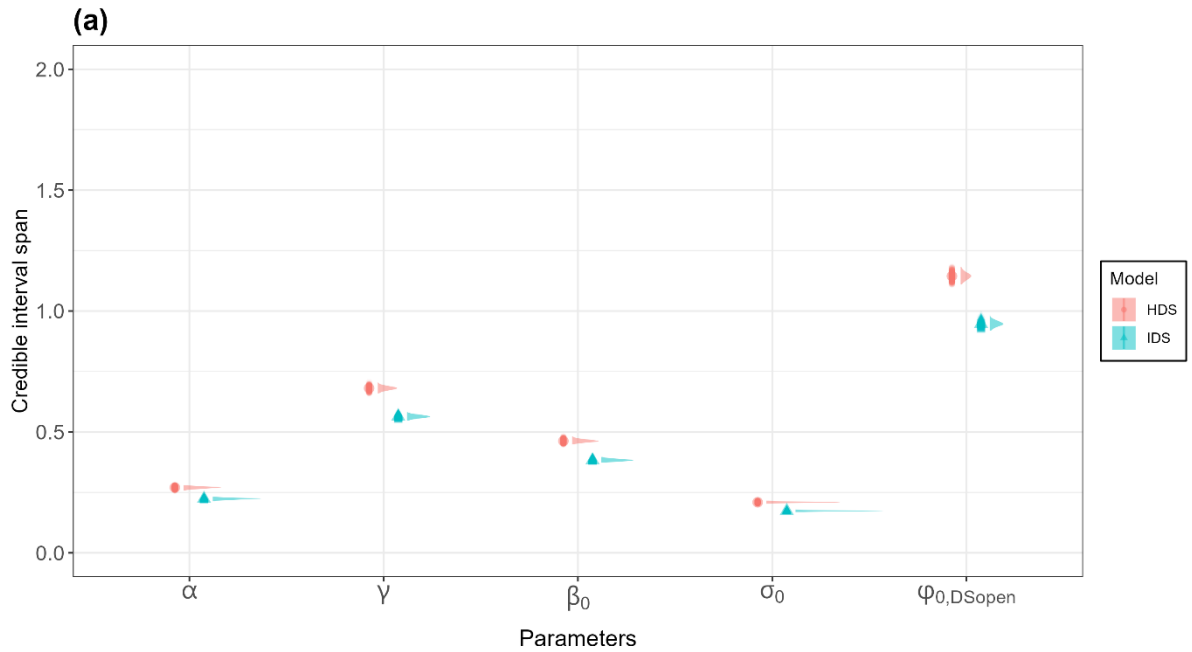
For each plot, we see in (a) average CRIs and their associated bootstrapped confidence intervals (CIs), depicted by vertical density plots, over simulated parameters. (b) CRIs average responses and CIs over simulated species availability ( $\varphi_{0,DSopen}$ ) and detectability ( $\sigma_0$ ) continuums on natural scales. Species availability averaged across multiple classes (0.05, 0.25, 0.5, 0.75 and 0.95 detection probability) are depicted by colour-graded lines. (c) CRIs average responses and CIs over multiple data collection cases. Proportions of semi-structured data (without temporal replicates) averaged over three classes (1, 3 and 6) are depicted by colour-graded lines. Each dot (a) and line (b-c) corresponds to a model response

from a bootstrapped resample of the considered subset. For comparison between HDS and IDS inferences accuracy, we plotted CRIs responses, grey ribbons, for the case of simulated species with high mean detectability ( $\sigma_0 = 200\text{m}$ ) and high probability of being available ( $\varphi_{0,DSopen} = 0.95$ ; B) surveyed over 100 sites with temporal replicates and six times the number of semi-structured sites (c), depicted with vertical lines. The Lower and upper bounds of the rectangles correspond to minimal and maximal estimated means CRIs values.

Bootstrap analysis | Number of data used : 500



Bootstrap analysis | Number of data used : 1500



#### IV.D References

- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bayraktarov E, Ehmke G, O'Connor J, Burns EL, Nguyen HA, McRae L, Possingham HP, Lindenmayer DB (2019) Do Big Unstructured Biodiversity Data Mean More Knowledge? *Frontiers in Ecology and Evolution*, 6.
- Besbeas P, Freeman SN, Morgan BJT, Catchpole EA (2002) Integrating Mark-Recapture-Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters. *Biometrics*, 58, 540–547.
- Blangiardo M, Cameletti M, Baio G, Rue H (2013) Spatial and spatio-temporal models with R-INLA. *Spatial and Spatio-temporal Epidemiology*, 7, 39–55. <https://doi.org/10.1016/j.sste.2013.07.003>
- Blondel J, Ferry C, Frochot B (1981) Point Counts with Unlimited Distance. *Studies in Avian Biology*.
- Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips T, Rosenberg KV, Shirk J (2009) Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience*, 59, 977–984. <https://doi.org/10.1525/bio.2009.59.11.9>
- Brlík V, Šilarová E, Škorpiťová J, Alonso H, Anton M, Aunins A, Benkő Z, Biver G, Busch M, Chodkiewicz T, Chylarecki P, Coombes D, de Carli E, del Moral JC, Derouaux A, Escandell V, Eskildsen DP, Fontaine B, Foppen RPB, Gamero A, Gregory RD, Harris S, Herrando S, Hristov I, Husby M, Ieronymidou C, Jiquet F, Kålås JA, Kamp J, Kmecl P, Kurlavičius P, Lehtikoinen A, Lewis L, Lindström Å, Manolopoulos A, Martí D, Massimino D, Moshøj C, Nellis R, Noble D, Paquet A, Paquet J-Y, Portolou D, Ramírez I, Redel C, Reif J, Ridzoň J, Schmid H, Seaman B, Silva L, Soldaat L, Spasov S, Staneva A, Szép T, Florenzano GT, Teufelbauer N, Trautmann S, van der Meij T, van Strien A, van Turnhout C, Vermeersch G, Vermouzek Z, Vikstrøm T, Voříšek P, Weiserbs A, Klvaňová A (2021) Long-term and large-scale multispecies dataset tracking population changes of common European breeding birds. *Scientific Data*, 8, 21. <https://doi.org/10.1038/s41597-021-00804-2>
- Buckland ST, Anderson DR, Burnham KP, Laake J, Borchers D, Thomas L (2001) *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford ; New York.
- Castagneyrol B, Valdés-Correcher E, Bourdin A, Barbaro L, Bouriaud O, Branco M, Centenaro G, Csóka G, Duduman M-L, Dulaurent A-M, Eötvös CB, Faticov M, Ferrante M, Fürjes-Mikó Á, Galmán A, Gossner MM, Harvey D, Howe AG, Kaennel-Dobbertin M, Koricheva J, Löveï GL, Lupaștean D, Milanović S, Mrazova A, Opgennoorth L, Pitkänen J-M, Popović M, Roslin TV, Scherer-Lorenzen M, Sam K, Tahadlová M, Thomas R, Tack AJM (2020) Can School Children Support Ecological Research? Lessons from the Oak Bodyguard Citizen Science Project. *Citizen Science: Theory and Practice*, 5. <https://doi.org/10.5334/cstp.267>
- Chandler RB, Royle JA, King DI (2011) Inference about density and temporary emigration in unmarked populations. *Ecology*, 92, 1429–1435. <https://doi.org/10.1890/10-2433.1>

- Cote C, Troncon C, Troncon C, Troncon C (2021) ROUTE 500® Version 3.0 - Descriptif de contenu. , 27.
- Devictor V, Whittaker RJ, Beltrame C (2010) Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions*, 16, 354–362. <https://doi.org/10.1111/j.1472-4642.2009.00615.x>
- Dietze MC (2017) Fusing Data Sources. In: *Ecological Forecasting* , pp. 118–130. Princeton University Press. <https://doi.org/10.2307/j.ctvc7796h.13>
- Donald PF, Fuller RJ (1998) Ornithological atlas data: a review of uses and limitations. *Bird Study*, 45, 129–145. <https://doi.org/10.1080/00063659809461086>
- Dorazio RM (2016) Bayesian data analysis in population ecology: motivations, methods, and benefits. *Population Ecology*, 58, 31–44. <https://doi.org/10.1007/s10144-015-0503-4>
- Doser JW, Finley AO, Weed AS, Zipkin EF (2021) Integrating automated acoustic vocalization data and point count surveys for estimation of bird abundance. *Methods in Ecology and Evolution*, 12, 1040–1049. <https://doi.org/10.1111/2041-210X.13578>
- Duchenne F, Porcher E, Mihoub J-B, Lois G, Fontaine C (2022) Controversy over the decline of arthropods: a matter of temporal baseline? *Peer Community Journal*, 2. <https://doi.org/10.24072/pcjournal.131>
- Farr MT, Zylstra ER, Ries L, Zipkin EF (2024) Overcoming data gaps using integrated models to estimate migratory species' dynamics during cryptic periods of the annual cycle. *Methods in Ecology and Evolution*, 15, 413–426. <https://doi.org/10.1111/2041-210X.14282>
- Fer I, Kelly R, Moorcroft PR, Richardson AD, Cowdery EM, Dietze MC (2018) Linking big models to big data: efficient ecosystem model calibration through Bayesian model emulation. *Biogeosciences*, 15, 5801–5830. <https://doi.org/10.5194/bg-15-5801-2018>
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315. <https://doi.org/10.1002/joc.5086>
- Fletcher RJ, Hefley TJ, Robertson EP, Zuckerberg B, McCleery RA, Dorazio RM (2019) A practical guide for combining data to model species distributions. *Ecology*, 100, e02710. <https://doi.org/10.1002/ecy.2710>
- Fox J, Monette G (1992) Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, 87, 178–183. <https://doi.org/10.1080/01621459.1992.10475190>
- Fox J, Weisberg S, Price B, Adler D, Bates D, Baud-Bovy G, Bolker B, Ellison S, Firth D, Friendly M, Gorjanc G, Graves S, Heiberger R, Krivitsky P, Laboissiere R, Maechler M, Monette G, Murdoch D, Nilsson H, Ogle D, Ripley B, Short T, Venables W, Walker S, Winsemius D, Zeileis A, R-Core (2023) *car: Companion to Applied Regression*.
- Gimenez O, Rossi V, Choquet R, Dehais C, Doris B, Varella H, Vila J-P, Pradel R (2007) State-space modelling of data on marked individuals. *Ecological Modelling*, 206, 431–438. <https://doi.org/10.1016/j.ecolmodel.2007.03.040>
- Guetté A, Carruthers-Jones J, Carver SJ (2021) *Projet CARTNAT Cartographie de la Naturalité*.
- Guillera-Arroita G (2017) Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography*, 40, 281–295. <https://doi.org/10.1111/ecog.02445>

- Guillera-Arroita G, Lahoz-Monfort JJ (2012) Designing studies to detect differences in species occupancy: power analysis under imperfect detection. *Methods in Ecology and Evolution*, 3, 860–869. <https://doi.org/10.1111/j.2041-210X.2012.00225.x>
- von Hirschheydt G, Stofer S, Kéry M (2023) “Mixed” occupancy designs: When do additional single-visit data improve the inferences from standard multi-visit models? *Basic and Applied Ecology*, 67, 61–69. <https://doi.org/10.1016/j.baae.2023.01.003>
- Hochachka WM, Fink D, Hutchinson RA, Sheldon D, Wong W-K, Kelling S (2012) Data-intensive science applied to broad-scale citizen science. *Trends in Ecology & Evolution*, 27, 130–137. <https://doi.org/10.1016/j.tree.2011.11.006>
- IGN (2011) Sylvoécorégions - Cartographie des sylvoécorégions. <https://geo.data.gouv.fr/fr/datasets/a40c533b984bdcd33d8a38f2430a117672395bc0>
- Isaac NJB, Jarzyna MA, Keil P, Dambly LI, Boersch-Supan PH, Browning E, Freeman SN, Golding N, Guillera-Arroita G, Henrys PA, Jarvis S, Lahoz-Monfort J, Pagel J, Pescott OL, Schmucki R, Simmonds EG, O’Hara RB (2020) Data Integration for Large-Scale Models of Species Distributions. *Trends in Ecology & Evolution*, 35, 56–67. <https://doi.org/10.1016/j.tree.2019.08.006>
- Issa N, Muller Y (2015) Atlas des oiseaux de France métropolitaine: Nidification et présence hivernale. DELACHAUX, Paris.
- Johnston A, Hochachka WM, Strimas-Mackey ME, Ruiz Gutierrez V, Robinson OJ, Miller ET, Auer T, Kelling ST, Fink D (2021) Analytical guidelines to increase the value of community science data: An example using eBird data to estimate species distributions. *Diversity and Distributions*, 27, 1265–1277. <https://doi.org/10.1111/ddi.13271>
- Johnston A, Matechou E, Dennis EB (2022) Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*, 14, 103–116. <https://doi.org/10.1111/2041-210X.13834>
- Johnston A, Moran N, Musgrove A, Fink D, Baillie SR (2020) Estimating species distributions from spatially biased citizen science data. *Ecological Modelling*, 422, 108927. <https://doi.org/10.1016/j.ecolmodel.2019.108927>
- Juat N, Meredith M, Kruschke J (2022) mcmcOutput: Functions to Store, Manipulate and Display Markov Chain Monte Carlo (MCMC) Output.
- Keil P, Wilson AM, Jetz W (2014) Uncertainty, priors, autocorrelation and disparate data in downscaling of species distributions. *Diversity and Distributions*, 20, 797–812. <https://doi.org/10.1111/ddi.12199>
- Kelling S, Johnston A, Bonn A, Fink D, Ruiz-Gutierrez V, Bonney R, Fernandez M, Hochachka WM, Julliard R, Kraemer R, Guralnick R (2019) Using Semistructured Surveys to Improve Citizen Science Data for Monitoring Biodiversity. *BioScience*, 69, 170–179. <https://doi.org/10.1093/biosci/biz010>
- Kellner K, Meredith M (2021) jagsUI: A Wrapper Around “rjags” to Streamline “JAGS” Analyses.
- Kellner KF, Swihart RK (2014) Accounting for Imperfect Detection in Ecology: A Quantitative Review. *PLOS ONE*, 9, e111436. <https://doi.org/10.1371/journal.pone.0111436>

- Kéry M, Royle JA (2015) *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models*. Academic Press, Amsterdam ; Boston.
- Kéry M, Royle JA (2020) *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS: Volume 2: Dynamic and Advanced Models*. Academic Press, Cambridge.
- Kéry M, Royle JA, Hallman T, Robinson WD, Strebel N, Kellner KF (2024) Integrated distance sampling models for simple point counts. *Ecology*, n/a, e4292. <https://doi.org/10.1002/ecy.4292>
- Kéry M, Royle A, Meredith M (2023) AHMbook: Functions and Data for the Book “Applied Hierarchical Modeling in Ecology” Vols 1 and 2.
- King R (2014) Statistical Ecology. *Annual Review of Statistics and Its Application*, 1, 401–426. <https://doi.org/10.1146/annurev-statistics-022513-115633>
- Lauret V, Labach H, Authier M, Gimenez O (2021) Using single visits into integrated occupancy models to make the most of existing monitoring programs. *Ecology*, 102, e03535. <https://doi.org/10.1002/ecy.3535>
- Lehikoinen A (2013) Climate change, phenology and species detectability in a monitoring scheme. *Population Ecology*, 55, 315–323. <https://doi.org/10.1007/s10144-012-0359-9>
- Lenth RV, Bolker B, Buerkner P, Giné-Vázquez I, Herve M, Jung M, Love J, Miguez F, Riebl H, Singmann H (2024) emmeans: Estimated Marginal Means, aka Least-Squares Means.
- MacKenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA, Langtimm CA (2002) Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One. *Ecology*, 83, 2248–2255. <https://doi.org/10.2307/3072056>
- Mackenzie DI, Royle JA (2005) Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology*, 42, 1105–1114. <https://doi.org/10.1111/j.1365-2664.2005.01098.x>
- Martin ME, Green DS, Garrison D, Hartman J, Jackson BK, Mackey H, McDonald MA, Smith BH, Smith TR, Vanderzwan SL, Stock SL, Sacks BN, Matthews SM (2023) An integrated spatial capture–recapture approach reveals the distribution of a cryptic carnivore in a protected area. *Ecosphere*, 14, e4634. <https://doi.org/10.1002/ecs2.4634>
- Meehan TD, Saunders SP, DeLuca WV, Michel NL, Grand J, Deppe JL, Jimenez MF, Knight EJ, Seavy NE, Smith MA, Taylor L, Witko C, Akresh ME, Barber DR, Bayne EM, Beasley JC, Belant JL, Bierregaard RO, Bildstein KL, Boves TJ, Brzorad JN, Campbell SP, Celis-Murillo A, Cooke HA, Domenech R, Goodrich L, Gow EA, Haines A, Hallworth MT, Hill JM, Holland AE, Jennings S, Kays R, King DT, Mackenzie SA, Marra PP, McCabe RA, McFarland KP, McGrady MJ, Melcer Jr R, Norris DR, Norvell RE, Rhodes Jr OE, Rimmer CC, Scarpignato AL, Shreading A, Watson JL, Wilsey CB (2022) Integrating data types to estimate spatial patterns of avian migration across the Western Hemisphere. *Ecological Applications*, 32, e2679. <https://doi.org/10.1002/eap.2679>
- Miller DL, Rexstad E, Thomas L, Marshall L, Laake JL (2019) Distance Sampling in R. *Journal of Statistical Software*, 89, 1–28. <https://doi.org/10.18637/jss.v089.i01>

- Mizel JD, Schmidt JH, Lindberg MS (2018) Accommodating temporary emigration in spatial distance sampling models. *Journal of Applied Ecology*, 55, 1456–1464. <https://doi.org/10.1111/1365-2664.13053>
- Nabias J, Barbaro L, Fontaine B, Dupuy J, Couzi L, Vallé C, Lorrilliere R (2024) Reassessment of French breeding bird population sizes using citizen science and accounting for species detectability.
- Navarro LM, Fernández N, Guerra CA, Rob Guralnick, Guralnick RP, Kissling WD, Londoño MC, Muller-Karger FE, Turak E, Balvanera P, Costello MJ, Delavaud A, Serafy GYE, Ferrier S, Geijzendorffer IR, Geller GN, Jetz W, Kim ES, Kim H, Martin CS, McGeoch MA, Mwampamba TH, Nel JL, Nicholson E, Pettorelli N, Schaeppman ME, Skidmore AK, Sousa-Pinto I, Pinto IS, Vergara SG, Vihervaara P, Xu H, Yahara T, Gill MJ, Pereira HM (2017) Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability*, 29, 158–169. <https://doi.org/10.1016/j.cosust.2018.02.005>
- Nichols JD, Thomas L, Conn PB (2009) Inferences About Landbird Abundance from Count Data: Recent Advances and Future Directions. In: *Modeling Demographic Processes In Marked Populations Environmental and Ecological Statistics*. (eds Thomson DL, Cooch EG, Conroy MJ), pp. 201–235. Springer US, Boston, MA. [https://doi.org/10.1007/978-0-387-78151-8\\_9](https://doi.org/10.1007/978-0-387-78151-8_9)
- Ogle K, Barber JJ (2020) Ensuring identifiability in hierarchical mixed effects Bayesian models. *Ecological Applications*, 30, e02159. <https://doi.org/10.1002/eap.2159>
- Pavanato Julião H (2021) Development of integrated distance sampling models. Thesis. University of Otago.
- Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Heip C, Höft R, Hurtt G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M (2013) Essential Biodiversity Variables. *Science*, 339, 277–278. <https://doi.org/10.1126/science.1229931>
- Plummer M (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*.
- Powney GD, Preston CD, Purvis A, Van Landuyt W, Roy DB (2014) Can trait-based analyses of changes in species distribution be transferred to new geographic areas? *Global Ecology and Biogeography*, 23, 1009–1018. <https://doi.org/10.1111/geb.12189>
- Proença V, Martin LJ, Pereira HM, Fernandez M, McRae L, Belnap J, Böhm M, Brummitt N, García-Moreno J, Gregory RD, Honrado JP, Jürgens N, Opige M, Schmeller DS, Tiago P, van Swaay CAM (2017) Global biodiversity monitoring: From data sources to Essential Biodiversity Variables. *Biological Conservation*, 213, 256–263. <https://doi.org/10.1016/j.biocon.2016.07.014>
- Roché J-E, Muller Y, Sibley J-P (2013) Une méthode simple pour estimer les populations d'oiseaux communs nicheurs en France. *Alauda*, 81, 241–268.
- Royle JA, Dorazio RM (2008) Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities.

- Rue H, Riebler A, Sørbye SH, Illian JB, Simpson DP, Lindgren FK (2017) Bayesian Computing with INLA: A Review. *Annual Review of Statistics and Its Application*, 4, 395–421. <https://doi.org/10.1146/annurev-statistics-060116-054045>
- Schaub M (2020) 9. Combining Counts of Unmarked Individuals and Demographic Data Using Integrated Population Models. In: *Population Ecology in Practice* (eds Murray DL, Sandercock BK), pp. 215–236. Wiley-Blackwell, Hoboken.
- Schaub M, Abadi F (2011) Integrated population models: a novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, 152, 227–237. <https://doi.org/10.1007/s10336-010-0632-7>
- Silva del Pozo M, Body G, Rerig G, Basille M (2023) Guide on harmonising biodiversity monitoring protocols across scales. *Biodiversa+*.
- Sollmann R, Gardner B, Williams KA, Gilbert AT, Veit RR (2016) A hierarchical distance sampling model to estimate abundance and covariate associations of species and communities. *Methods in Ecology and Evolution*, 7, 529–537. <https://doi.org/10.1111/2041-210X.12518>
- Soroye P, Ahmed N, Kerr JT (2018) Opportunistic citizen science data transform understanding of species distributions, phenology, and diversity gradients for global change research. *Global Change Biology*, 24, 5281–5291. <https://doi.org/10.1111/gcb.14358>
- Sullivan BL, Wood CL, Iliff MJ, Bonney RE, Fink D, Kelling S (2009) eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142, 2282–2292. <https://doi.org/10.1016/j.biocon.2009.05.006>
- Thierion V, Vincent A, Valero S (2022) Theia OSO Land Cover Map 2020. <https://doi.org/10.5281/zenodo.6538861>
- Thompson WL (2002) Towards Reliable Bird Surveys: Accounting for Individuals Present but not Detected. *The Auk*, 119, 18–25. [https://doi.org/10.1642/0004-8038\(2002\)119\[0018:TRBSAF\]2.0.CO;2](https://doi.org/10.1642/0004-8038(2002)119[0018:TRBSAF]2.0.CO;2)
- Twining JP, Fuller AK, Sun CC, Calderón-Acevedo CA, Schlesinger MD, Berger M, Kramer D, Frair JL (2024) Integrating presence-only and detection/non-detection data to estimate distributions and expected abundance of difficult-to-monitor species on a landscape-scale. *Journal of Applied Ecology*, n/a. <https://doi.org/10.1111/1365-2664.14633>
- Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution*, 16, 446–453. [https://doi.org/10.1016/S0169-5347\(01\)02205-4](https://doi.org/10.1016/S0169-5347(01)02205-4)
- Zipkin EF, Inouye BD, Beissinger SR (2019) Innovations in data integration for modeling populations. *Ecology*, 100, 1–3.
- Zipkin EF, Zylstra ER, Wright AD, Saunders SP, Finley AO, Dietze MC, Itter MS, Tingley MW (2021) Addressing data integration challenges to link ecological processes across scales. *Frontiers in Ecology and the Environment*, 19, 30–38. <https://doi.org/10.1002/fee.2290>

## V Chapitre 3 : Effet de la non prise en compte du processus d'observation dans l'estimation d'indice de biodiversité

---

### V.A Contexte

Même si l'abondance d'une espèce est un indice clé pour l'étude et le rapportage (Pereira et al., 2013), elle reste insuffisante pour appréhender pleinement la complexité de la biodiversité. L'étude des communautés se définit par la détermination des dynamiques décrivant la diversité, *i.e.* pluralité taxonomique, phylogénétique et fonctionnelle des espèces coexistant dans le temps et dans l'espace (Loreau, 2012). La majorité des indices utilisent l'abondance relative (au sens de la proportion d'individus d'une espèce par rapport au nombre total d'individus) comme information supplémentaire afin de renseigner sur d'autres facettes des communautés (Maurer & McGill, 2010). Cependant, chaque espèce possède des caractéristiques morphologiques et comportementales spécifiques (Figure 27; Sóllymos et al., 2018). Ignorer ces différences pourrait biaiser considérablement les estimations des indices de communauté (Jarzyna & Jetz, 2016).

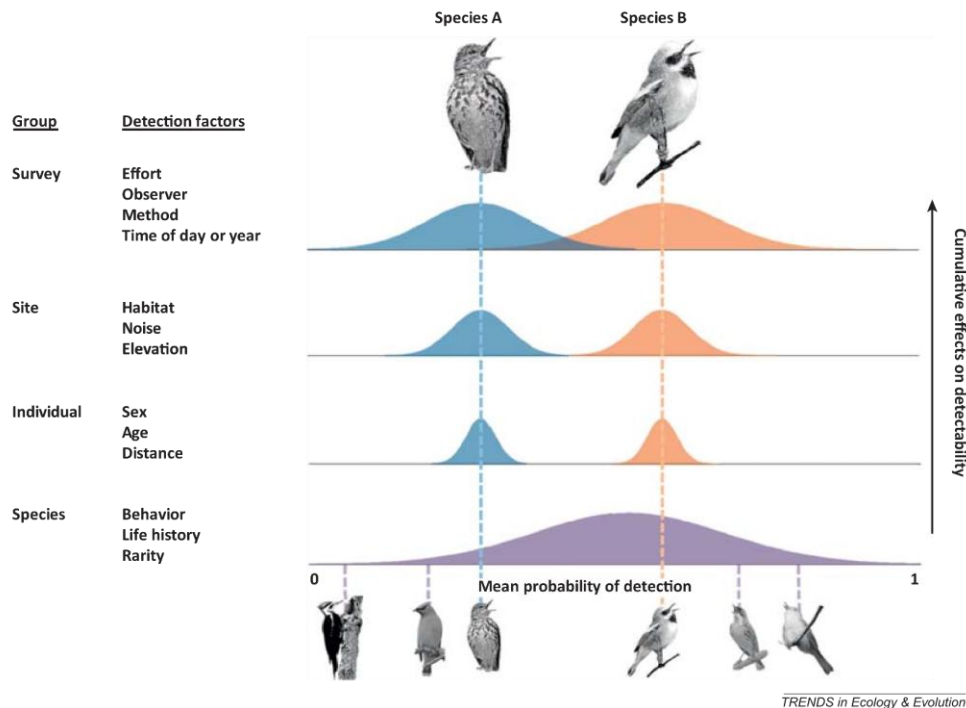


Figure 27 : Ensemble des éléments affectant la probabilité de détection des espèces. Illustration dans le cadre de modèles hiérarchiques de communauté. On peut considérer la détectabilité moyenne comme étant un trait spécifique décrit par le comportement, les traits d’histoires de vie et la rareté des espèces. La figure représente la distribution des probabilités de détection de deux oiseaux issus d’une même communauté selon différents éléments liés au processus d’observation (individuel, spécifiquement lié aux sites et aux visites). Figure tirée de [Iknayan et al., 2014](#).

Au sein des précédents chapitres, nous avons précisé dans quelle mesure les estimations d’abondance pouvaient être négativement biaisées lorsque le processus d’observation n’était pas pris en compte. Ici, nous comptons étudier les potentielles divergences de plusieurs indices de communautés, issues de l’omission du processus d’observation. L’étude se base sur des données du STOC-EPS qui a servi à plusieurs reprises au développement et à l’estimation d’indices de communautés ([Jiguet et al., 2005](#); [Julliard et al., 2006](#); [Princé, Moussus & Jiguet, 2012](#); [Godet et al., 2015](#)). Comme pour le précédent chapitre, nous avons réalisé une étude de simulation et une étude de cas. L’étude de simulation nous a permis de vérifier que les modèles développés parvenaient à correctement estimer les paramètres des processus écologiques et d’observation d’un ensemble d’espèces.

Nous avons ensuite utilisé les modèles sur les données du STOC-EPS entre 2015 et 2023 sur une cinquantaine d'espèces d'oiseaux, principalement communes.

Les résultats de ce chapitre montrent que, dans l'ensemble, les indices de communautés calculés à partir d'abondance ne prenant pas en compte le processus d'observation sont biaisés. L'étude de simulation détermine que cet écart est plus important lorsque l'indice de communauté (typiquement fonctionnel, comme par exemple l'indice de spécialisation à l'habitat) est corrélé à la détectabilité des espèces. La comparaison des indices de communautés estimés en corrigeant, ou non, le processus d'observation nous montre un biais global dans l'estimation des compositions taxonomique, phylogénétique et fonctionnelle des communautés.

### **Travail en cours**

Les résultats et conclusions du prochain chapitre sont issus d'une recherche encore en cours. La lecture de ce chapitre doit se faire en tenant compte du fait que les points suivants sont envisagés pour la version finale :

- Il est prévu de faire passer le nombre de communautés dans la simulation de 60 à 300. De plus, le nombre d'espèces par communauté passera de 10 à 15, dans le but de rendre les communautés plus comparables à celles observées par le STOC-EPS.
- Le nombre d'itérations choisi pour les MCMC des études de simulation et l'étude de cas, n'est pas suffisant. Les résultats présentés proviennent de modèles n'ayant pas convergé. On compte augmenter ce nombre d'itérations.
- L'étude de cas a été réalisée sur un sous-ensemble des données disponibles (*i.e.* tirage aléatoire de 100 carrés STOC sur l'ensemble de la période temporelle). On compte augmenter le nombre de sites, et donc affiner le gradient environnemental.
- Quatre espèces, à savoir, l'hirondelle rustique (*Hirundo rustica*), le martinet noir (*Apus apus*), la buse variable (*Buteo buteo*) et le faucon crécerelle (*Falco tinnunculus*) ont été exclus de l'analyse de cette article et (voir les espèces grisées dans l'appendice S3 de : *Supplementary information*) pourraient être à l'origine de biais dans l'un des indices de communautés (Community Dispersal Index; CDI) qui estime les capacités de dispersion des espèces. Cet indice se base sur la mesure de la surface d'aile qui est importante sur ces espèces. Bien que l'exclusion *ad hoc* de ces espèces n'ait pas altéré les résultats principaux, nous prévoyons de les retirer de l'étude finale pour éviter les biais potentiels liés au calcul de l'abondance relative.

**Référence du contexte**

- Godet L, Devictor V, Burel F, Robin J-G, Ménanteau L, Fournier J. 2016. Extreme Landscapes Decrease Taxonomic and Functional Bird Diversity but Promote the Presence of Rare Species. *Acta Ornithologica* 51:23–38. DOI: 10.3161/00016454AO2016.51.1.003.
- Godet L, Gaiüzere P, Jiguet F, Devictor V. 2015. Dissociating several forms of commonness in birds sheds new light on biotic homogenization. *Global Ecology and Biogeography* 24:416–426. DOI: 10.1111/geb.12266.
- Jarzyna MA, Jetz W. 2016. Detecting the Multiple Facets of Biodiversity. *Trends in Ecology & Evolution* 31:527–538. DOI: 10.1016/j.tree.2016.04.002.
- Jiguet F, Julliard R, Couvet D, Petiau A. 2005. Modeling Spatial Trends in Estimated Species Richness using Breeding Bird Survey Data: A Valuable Tool in Biodiversity Assessment. *Biodiversity & Conservation* 14:3305–3324. DOI: 10.1007/s10531-004-0448-y.
- Julliard R, Clavel J, Devictor V, Jiguet F, Couvet D. 2006. Spatial segregation of specialists and generalists in bird communities. *Ecology Letters* 9:1237–1244. DOI: 10.1111/j.1461-0248.2006.00977.x.
- Maurer BA, McGill BJ. 2010. Measurement of species diversity. In: Magurran AE, McGill BJ eds. *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford, New York: Oxford University Press, 55–64.
- Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Heip C, Höft R, Hurtt G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M. 2013. Essential Biodiversity Variables. *Science* 339:277–278. DOI: 10.1126/science.1229931.
- Princé K, Moussus J-P, Jiguet F. 2012. Mixed effectiveness of French agri-environment schemes for nationwide farmland bird conservation. *Agriculture, Ecosystems & Environment* 149:74–79. DOI: 10.1016/j.agee.2011.11.021.
- Sólymos P, Matsuoka SM, Stralberg D, Barker NKS, Bayne EM. 2018. Phylogeny and species traits predict bird detectability. *Ecography* 41:1595–1603. DOI: 10.1111/ecog.03415.

V.B Article

# **Accounting for species detectability improves bird taxonomic, functional and phylogenetic community metrics**

*En préparation*

Jean Nabias<sup>1,2,\*</sup>, Romain Lorrillière<sup>2,3</sup>, Emmanuelle Porcher<sup>2</sup>, Clément Vallé<sup>2</sup>, Benoit Fontaine<sup>2,4</sup>, Luc Barbaro<sup>2,5</sup>

1: LPO-BirdLife France, Fonderies Royales, Rochefort Cedex, France

2: CESCO, Muséum National d'Histoire Naturelle, CNRS, Sorbonne-University, Paris, France

3: Centre de Recherches sur la Biologie des Populations d'Oiseaux (CRBPO), MNHN-CNRS-OFB, Paris, France

4: PATRINAT, OFB, Muséum National d'Histoire Naturelle, CNRS, Paris, France

5: Dynafor, INRA-INPT, University of Toulouse, Auzeville, France

Keywords : Breeding Bird Survey ; Hierarchical Model ; Bird Community ; Observation process ; Community diversity

## Abstract

Because of the unprecedented scale of global change effects on biodiversity, a detailed understanding of species community composition, either taxonomic, phylogenetic or functional, is essential for implementing effective conservation measures.

Species abundance represents an essential variable informing species conservation status but also provides crucial insights into community composition. However, the number of individuals detected during sampling does not represent a constant proportion of their total population size due to variations induced by habitat- or effort-related factors during data collection, *i.e.* observation process. Specific life-history traits of species within communities, which define their morphology and behaviour, can impact their detectability and could potentially result in inaccurate representations of community composition and function.

In this study, we used data from the French Breeding Bird Survey (FBBS), to assess whether the omission of the observation process could lead to misrepresentations of community composition. We computed two models (MRelat and MHDS), respectively, omitting and accounting for the observation process while estimating species abundance. We first conducted a simulation study to assess community indices, more specifically averaged traits-based indices, *i.e.* Community Weighted Means (CWMs), across a gradient of correlation between the targeted trait and species detectability affecting the number of individuals detected during sampling occasions. We complemented this study with field data collected by the FBBS over 100 sites during 2015-2023 breeding seasons examining multiple CWM, as well as indices depicting phylogenetic and taxonomic diversity. Results from our simulation study showed significant biases in estimated community indices, when species traits were moderately correlated to species detectability ( $r \geq 0.5$ ). These results were corroborated by our case study, which revealed substantial differences in taxonomic,

phylogenetic and functional community indices between the two models across various French habitats, highlighting the risk of misrepresenting community composition if the observation process is omitted.

## Introduction

Ecological community, as the assemblage of multiple interacting species at a given space and time (Begon et al., 2006), can be described by different metrics based on species richness (number of distinct species), evenness (species abundance disparities among community) or a combination of both (Maurer and McGill, 2010). While these metrics provide a global description of community composition, they tend to overlook the crucial role of functional diversity, *i.e.* the value and range of value of species traits affecting ecosystem functions (Tilman, 2001). Assessment of functional diversity requires either the depiction of the functional-trait space of the community (Villéger et al., 2008) or the use of surrogates such as Community Weighted Mean traits (CWMs) that average species trait values to their relative abundance in communities (Violle et al., 2007).

Community indices are essential to track substantial changes in biodiversity (Pilotto et al., 2020) and monitoring biotic homogenization of ecosystems (Blowes et al., 2024; Lockwood and McKinney, 2001; Olden, 2006). These changes arise from introduction of invasive species (Mathieu et al., 2024; Maurer et al., 2001), species thermal tolerance to climate change (Lehikoinen et al., 2021; Princé and Zuckerman, 2015) or the decline of specialists (Clavel et al., 2011; Devictor et al., 2008b). These indices also play a crucial role in informing conservation action, notably through assessment of ecosystem functions (Ali, 2023; Cadotte et al., 2011). However, their reliance on sample-based collections of species relative abundance in communities (Maurer and McGill, 2010) could produce biased estimates when omitting the observation process (Roth et al., 2018).

In contrast to studies based on experiments in controlled environments (Tilman et al., 2001), field studies of species communities often rely on detection-non detection or transect counts to estimate the presence or quantity of individuals of each targeted species across sampled locations. Despite efforts to collect representative data over a large range of sampled habitats, the observation process is still often marginally taken into account in analyses (Bonar et al., 2010; Iknayan et al., 2014; Jarzyna and Jetz, 2016). Thus, most field surveys either imply that collected data are a complete census (*i.e.* considering that all individuals of the targeted population are detected) or assume that the sampled individuals always correspond to the same proportion of the wild population (Ross and Reeve, 2011; Thompson, 2002). Yet, omitting the observation process accounting for sampling conditions, either habitat- or effort-related, is now widely acknowledged as a main source of ecological and statistical bias (Kellner and Swihart, 2014; MacKenzie et al., 2002). This bias is even more consequential for the estimation of community indices (Jarzyna and Jetz, 2016), where species detection probabilities can be tied to functional traits, such as habitat specialisation (Johnston et al., 2014), leaf area and seed mass (Roth et al., 2018); species abundance (Morelli et al., 2022); or can be phylogenetically related (Chen et al., 2013; Sóllymos et al., 2018), leading to misrepresentation of community response (Iknayan et al., 2014; Palacio et al., 2020; Tingley and Beissinger, 2013).

Accounting for the observation process requires specific data collection, such as temporal replicates, capture-recapture, distance sampling (King, 2014), but also specific statistical framework, such as Hierarchical Models (HM), allowing distinct modelling of both observation and ecological process and ensued community indices (Dorazio et al., 2006; Kéry et al., 2009). As we aimed to assess the importance of the observation process for estimation of species community indices, we based our study on two hypotheses. **(H1)** We expect that community indices are biased when the observation process is omitted. **(H2)** We expect

greater biases towards community metrics based on species traits and functions, especially CWMs, because detectability varies among trait values or categories. Given **H1** and **H2**, we expect that these biases also occur for other widely used community metrics, such as taxonomic and phylogenetic diversity (**H3**).

To test our hypothesis, we relied on a simulation and case study based on robust design schemes (Pollock, 1982) with data collection of observation distances at given sites with within- and between-year replicates. This allows calibrating Hierarchical Distance Sampling (HDS) models accounting for temporary emigration (Chandler et al., 2011), *i.e.* species probabilities of being exposed to sampling events. For **H1**, we conducted a simulation study comparing CWM estimated from a model omitting the observation process (**M<sub>Relat</sub>**) and a model accounting for it (**M<sub>HDS</sub>**) using the same generated data. To test **H2**, we examined the influence of species traits-to-species detectability correlation by simulating multiple scenarios of correlation based on generated species detectability. We investigated **H2** and **H3** using real-life citizen science data collected by the French Breeding Bird Survey (FBBS; Jiguet et al., 2012) from 2015 to 2023 breeding season and computed four CWMs accounting for body mass, dispersal, specialisation and thermal preferences. We hypothesised further that each CWM is differently affected by the modelled observation process, *i.e.* trait value detectability. Hence, we relied on a gradient of species traits, with presumed correlation to detectability, ranging from traits related to species morphometrics and ecology, as well as their phylogenetic and taxonomic specificities.

Material and methods

*Hierarchical Distance Sampling framework*

Hierarchical Distance Sampling model (HDS) corresponds to a state-space model allowing the distinct modelling of the abundance process, depicting abundance variations across space and time, while taking account of the observation process that determines data collection (Kéry and Royle, 2015).

In our case, we consider that the local species community at site  $s$  and year  $t$ , composed of  $p$  species follows a poisson distribution with  $M_{s,t,p}$ , the local abundance of specie  $p$  :

$$M_{s,t,p} \sim \text{Poisson}(\lambda_{s,t,p})$$

From this local abundance, given a sampling design with replicated visits  $k$ , the number of individual from the species  $p$  can be modelled as a binomial outcome with probability ( $\varphi_{k,p}$ ) of the individuals being considered as available for detection (for instance, birds vocalisation during observers lists) :

$$N_{s,k,t,p} \sim \text{Binomial}(M_{s,t,p}, \varphi_{k,p})$$

For Distance sampling approaches, observers are tasked to collect observation distance (Buckland et al., 2015). Thus allowing the determination of a detection function  $f$  of parameter ( $\sigma_{s,k,t,p}$ ) depicting the probability of detection of available individuals ( $N_{s,k,t,p}$ ) given the observation distance ( $x_{s,k,t,p}$ ). A vector of cell probabilities ( $\pi_{s,k,t,p}$ ) is derived from the detection function  $f$  with each cell corresponding to probabilities assigned to distance bins intervals (Kéry and Royle, 2015).

$$y_{s,k,t,p} \sim \text{Multinomial}(N_{s,k,t,p}, \pi_{s,k,t,p}), \text{ with } \pi_{s,k,t,p} = f(x_{s,k,t,p}, \sigma_{s,k,t,p})$$

### Simulation study

For the simulation study, we simulated 60 communities each of 10 species. We generated a set of species parameters depicting species abundance ( $\beta_{0,p}$ ), species abundance trend ( $\beta_p^{trend}$ ), species availability ( $\varphi_{0,p}$ ) and species detectability ( $\sigma_{0,p}$ ) as well as coefficients ( $\beta_p$ ,  $\gamma_p$  and  $\alpha_p$ ) over covariables ( $X_s$ ,  $U_s$ ,  $Z_{s,k,t}$ ) effects on abundance, availability and detectability collected at sites  $i$ , visits  $k$  and year  $t$  for species  $p$  (Eq. 1). For each community, we simulated a sampling design of 500 sites, visited twice over nine years. Ranges of simulated parameters were based upon previous studies Nabias et al., 2024.

$$\begin{cases} \log(\lambda_{s,t,p}) = \beta_{0,p} + \beta_p X_s + \beta_p^{trend} Year_t \\ \text{logit}(\varphi_{k,p}) = \varphi_{0,p} + \gamma_p U_k \\ \log(\sigma_{s,k,t,p}) = \sigma_{0,p} + \alpha_p Z_{s,k,t} \end{cases} \quad (\text{Eq 1})$$

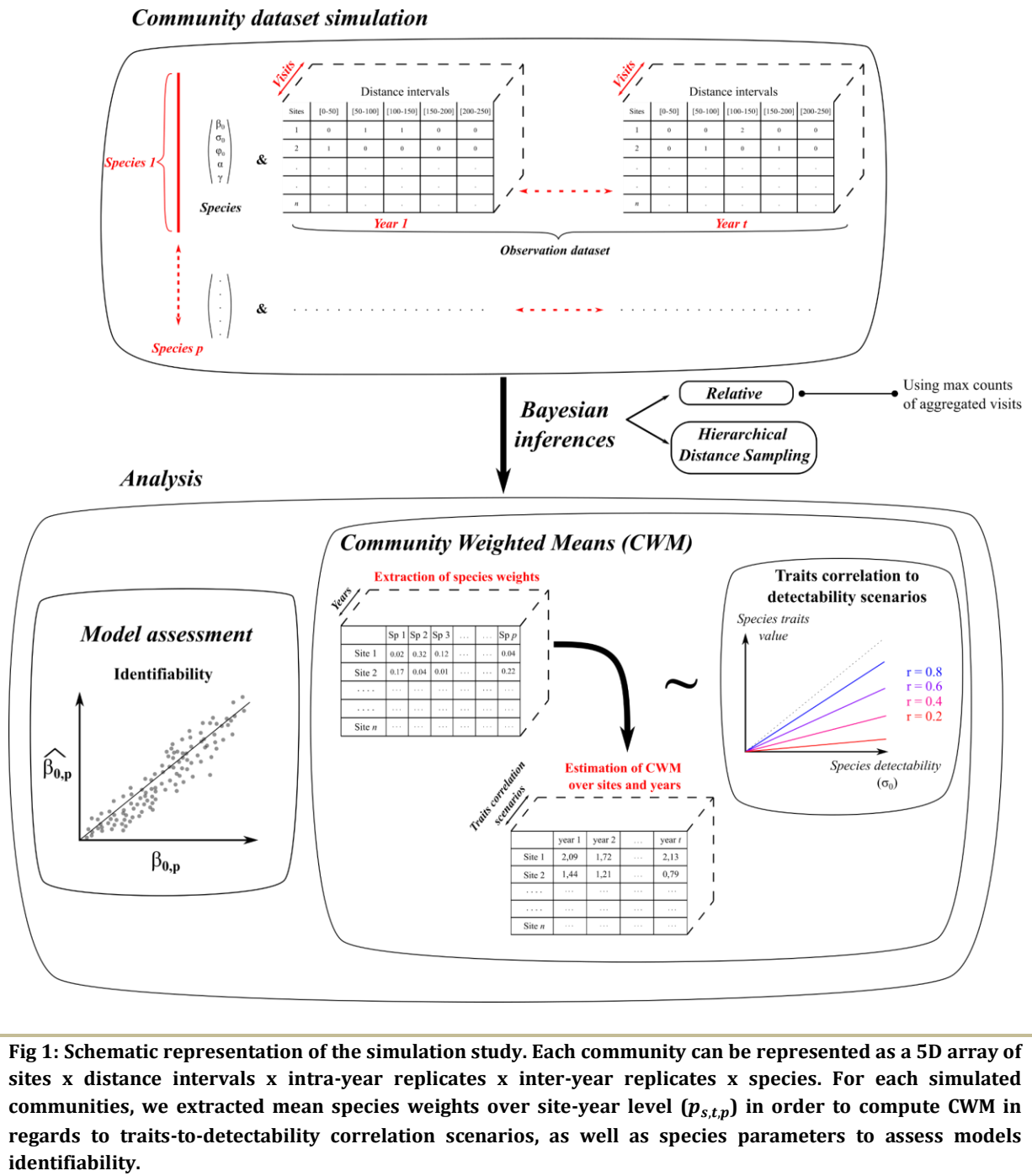
To test **H1** and **H2** in a simulated environment, we formulated two models (1) a relative model ( $\mathbf{M}_{\text{Relat}}$ ) that modelled species abundance and its temporal variation while disregarding the observation process and (2) a HDS model ( $\mathbf{M}_{\text{HDS}}$ ) that takes account of the observation process by estimating species availability ( $\varphi_{k,p}$ ) and detectability ( $\sigma_{s,k,t,p}$ ). We accounted for the observation process during data simulation of both models. However, for  $\mathbf{M}_{\text{Relat}}$ , we degrade observed data, only considering the maximal number of counted individuals (*i.e.* sum of counts over the distance intervals) over the visits at each site  $s$  and years  $t$  (Fig 1).

In order to compute Community Species Weights (CWM), we extracted estimated species abundances in communities ( $M_{s,t,p}$ ) to estimate their weights in community ( $p_{s,t,p}$ ) for the computation of community weights means ( $CWM_{s,t}$ ) at each site  $s$  over years  $t$ , see

*Diversity metrics* section for more details. As we expect potential variation of estimated abundance in regards to species trait correlation, but finding no prior studies simulating the correlation between species traits and detection, we computed  $CWM_{s,t}$  based upon a gradient of correlation scenarios  $r = \{0; 0.2; 0.4; 0.6; 0.8; 1\}$ , between species traits ( $\theta_p$ ) and species detectability ( $\sigma_{0,p}$ ). This allowed us to compare “true” CWM ( $CWM_{s,t}^{true}$ ) based upon simulated species parameter and correlation traits-to-detectability scenarios ( $r$ ) to estimated CWM ( $CWM_{s,t}$ ) from  $\mathbf{M}_{Relat}$  and  $\mathbf{M}_{HDS}$  using estimated species weights at the site-year level (Fig 1).

To assess if  $\mathbf{M}_{Relat}$  and  $\mathbf{M}_{HDS}$  could correctly estimate CWM, we modelled the difference ( $\Delta_{s,t} = CWM_{s,t}^{true} - CWM_{s,t}$ ), using a linear model considering the interaction between the model formulation (*i.e.* either Relative or HDS) and the correlation scenarios ( $r$ ). We included two random effects, one over the years, as we implied a temporal trend, and another over the community simulation indexes, as we generated twice the same community for each model formulation. As we calibrate model from simulation outputs, we assessed the signal robustness of  $\Delta_{s,t}$  response to covariates using bootstrap analysis. The bootstrap analysis consisted of calibrating the model over 100 iterations, with each iteration resampling 70% of the output simulated communities.

We assessed model identifiability, *i.e.* ability to correctly estimated parameters, through R-squared coefficient between simulated and estimated species parameters for both models (Fig 1). For data generation, we relied on a modified version of *simHDSopen* function from AHMBook (Kéry et al., 2023).  $\mathbf{M}_{Relat}$  and  $\mathbf{M}_{HDS}$  models were calibrated in JAGS 4.3.1 (Plummer, 2003) using jagsUI (Kellner and Meredith, 2021). See appendix S1 for used priors and Markov Chain Monte Carlo (MCMC) parameters.



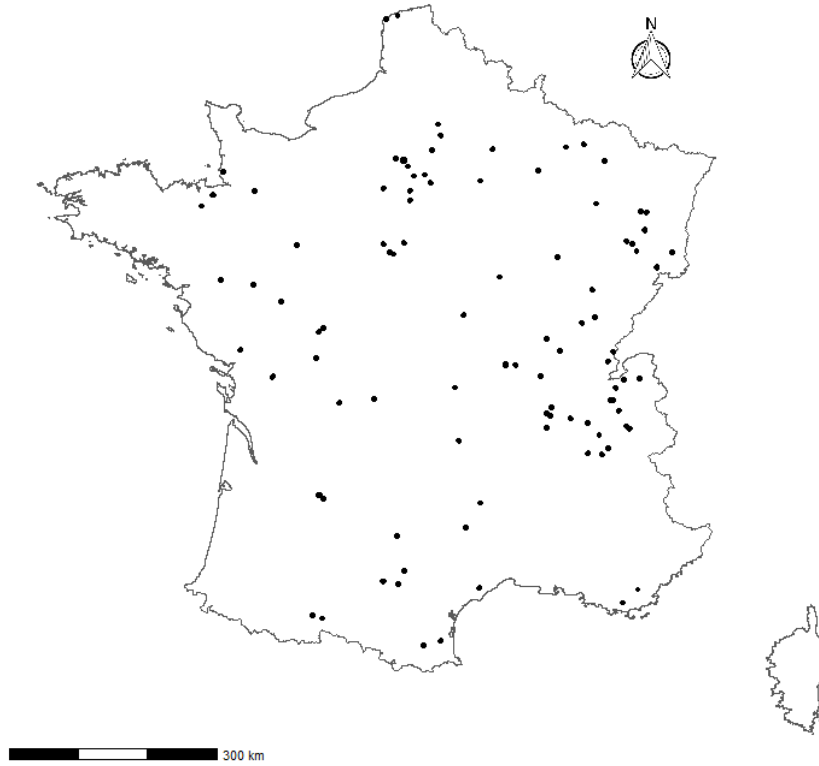
**Fig 1:** Schematic representation of the simulation study. Each community can be represented as a 5D array of sites x distance intervals x intra-year replicates x inter-year replicates x species. For each simulated communities, we extracted mean species weights over site-year level ( $p_{s,t,p}$ ) in order to compute CWM in regards to traits-to-detectability correlation scenarios, as well as species parameters to assess models identifiability.

*Case study*

*Bird data*

We studied 51 bird species using data from the French Breeding Bird Survey (FBBS; Jiguet et al., 2012), a citizen science scheme aiming to collect standardised data over France from skilled birders. It consists of a 5-min point count where birders are tasked to estimate the distance of observation of detected individuals, either visual or auditory contact, using distance classes varying from 0-25 ; 25-100 ; 100-200 and 200-more. The sampling design is based upon a systematic grid from which birders are assigned on a 2x2km grid upon their locality (less than 10km). Birders select ten spatial locations at least 300m apart, depicting habitat distribution of the 2x2km grid, which correspond to the spatial unit of our study. Data is collected over the breeding season, from March to June and birders are tasked to visit the locations two to three times each year.

For our study, we filter spatial location in order to keep locations that were followed constantly across nine years. We also decided to remove individuals detected “in flight” for the distance sampling analysis. In total, we used data collected over 1000 sites by 92 observers, assigned over 100 2x2km grids in France (Fig 2).



**Fig 2: Spatial distribution of FBBS sites in France. Each dot corresponds to the location of 2x2km grids (n = 100).**

### Model formulation

We modelled abundance and variation across sites ( $s$ ), years ( $t$ ) and species ( $p$ ) using a Poisson distribution of parameter ( $\lambda_{s,t,p}$ ). This model depicts species abundance variation across habitats ( $\beta_{h,p}^{cat}$ ) and years, considering linear trends over the year ( $\beta_p^{trend}$ ). Habitat corresponds to categorical covariates ( $H$ ) informed by observers. We also included a random effect over sites ( $\varepsilon_s^{abund}$ ) nested across 10x10km grids (See S1).

$$M_{s,t,p} \sim \text{Poisson}(\lambda_{s,t,p})$$

$$\log(\lambda_{s,t,p}) = \beta_{0,p} + \sum_{h=1}^H \beta_{h,p}^{cat} \text{Habitat}_s + \beta_p^{trend} \text{Year}_t + \varepsilon_s^{abund}$$

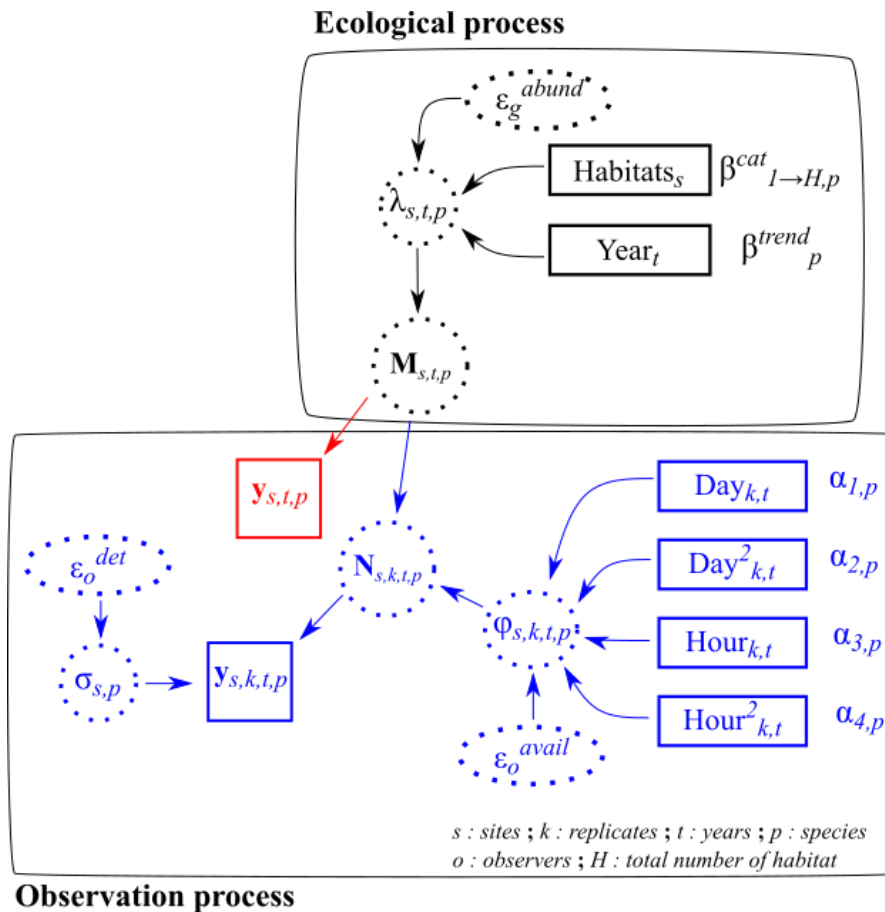
For  $\mathbf{M}_{\text{Relat}}$  model, we did not take account of the observation process (Fig 3), thus considering  $(y_{s,t,p})$  as the maximum number of counted individuals across intra-year replicates, *i.e.* being the sum over site-year-replicates distance bins intervals observations. For the  $\mathbf{M}_{\text{HDS}}$  model, we take account of the observation process  $(y_{s,k,t,p})$ , considering that individuals exposed to the sampling design  $(N_{s,k,t,p})$  could vary across visits during the breeding season and years, thus we decided to model species availability  $(\varphi_{s,k,t,p})$  as the quadratic effect of Julian day and hours from sunrise to capture birds' breeding phenology and dawn chorus (Fig 3).

$$\text{logit}(\varphi_{s,k,t,p}) = \varphi_{0,p} + \alpha_{1,p} \text{Day}_{k,t} + \alpha_{2,p} \text{Day}_{k,t}^2 + \alpha_{3,p} \text{HrSun}_{k,t} + \alpha_{4,p} \text{HrSun}_{k,t}^2 + \varepsilon_o^{\text{avail}}$$

We relied on a half-normal function of parameter  $(\sigma_{s,p})$  to model species detectability, the probability of detecting individuals considered available. As the distance classes collected by observers are not equivalent, with coarser intervals being farther from the observer, we modelled the cell probabilities  $(\pi_{s,p})$  using finer distance bins, of 25m each, which we summed in order to describe the distance bin from the sampling design (see appendix S2 for bias study over distance bins aggregation).

$$\log(\sigma_{s,p}) = \sigma_{0,p} + \varepsilon_o^{\text{det}}$$

As multiple observers participate in the survey, we added random effects  $(\varepsilon_o^{\text{avail}}, \varepsilon_o^{\text{det}})$  to take account of this source of variability.



**Fig 3 : Directed acyclic graph representing model structure of  $M_{HDS}$ . Solid rectangles correspond to observed data, while dotted lines indicate latent variables. Coefficients are represented on the side of their respective covariates. Both model formulations are depicted using colours. Coefficients, latent variables and observed data used in both formulations, solely for  $M_{Relat}$  and for  $M_{HDS}$  are represented in black, red and blue. For  $M_{Relat}$ , we did not account for covariate and latent variables of the observation process, instead we considered individual count at the site-year level ( $y_{s,t,p}$ ) by degrading observed data and using max count of individuals (i.e. sum of counts over the distance intervals) over the visits at each site and year.**

### Diversity Metrics

#### Community Weighted Means (CWM) metrics and species traits

Community Weighted Means (CWM) summarise the distribution of species traits ( $\theta_p$ ) at the community level (composed of  $P$  species in total), in our case at the site-year level, as the weighted sum of species traits in regards to their relative abundance in the community (Gaüzère et al., 2019):

$$CWM_{s,t} = \sum_{p=1}^P p_{s,t,p} \theta_p$$

In order to compute CWM at the site-year level ( $CWM_{s,t}$ ), we estimated species weights in community ( $p_{s,t,p}$ ) as their relative abundance to the sum of estimated abundance of over species at the site-year level :

$$p_{s,t,p} = \frac{M_{s,t,p}}{\sum_{p=1}^P M_{s,t,p}}$$

For species traits ( $\theta_p$ ), we used a gradient of species traits, ranging from species body mass; Species Specialisation Index to habitats (SSI; Julliard et al., 2006); Hand-Wing Index (HWI) an index of dispersal abilities (Arango et al., 2022; Claramunt et al., 2012); and Species Thermal Index (STI; Devictor et al., 2008). While body mass and SSI are known to affect species detectability (Johnston et al., 2014), we expect a slight-to-no correlation for HWI as it is not directly linked to wingspan known to affect detectability (Nolte et al., 2016). We did not have *a priori* expectation of correlation between STI and detectability. We retrieved species body mass and HWI from the AVONET database (Tobias et al., 2022), while SSI and STI were estimated from previously collected FBBS data. See Appendix S3 for species traits data.

### *Phylogenetic and taxonomic diversity*

For phylogenetic diversity, we used the Rao Quadratic entropy (RaoQ; Rao, 1982) to estimate site-year level phylogenetic heterogeneity (Rocchini et al., 2017). For the pairwise phylogenetic distances ( $d_{pv}$ ) between species  $p$  et  $v$ , we used the phylogeny assembled by Burleigh et al., 2015. For species relative abundance ( $p_{s,t,p}, p_{s,t,v}$ ), we relied on estimated species abundances obtained through  $\mathbf{M}_{\text{Relat}}$  and  $\mathbf{M}_{\text{HDS}}$ .

$$RaoQ_{s,t} = \sum_{p=1}^{P_{s,t}} \sum_{v=1}^{P_{s,t}} p_{s,t,p} p_{s,t,v} d_{pv}$$

For taxonomic diversity, we measured the Simpson diversity index, which correspond to the estimation of the RaoQ without accounting for phylogenetic relatedness, implying  $d_{pv} = 1$  for each species pair (Ricotta, 2005).

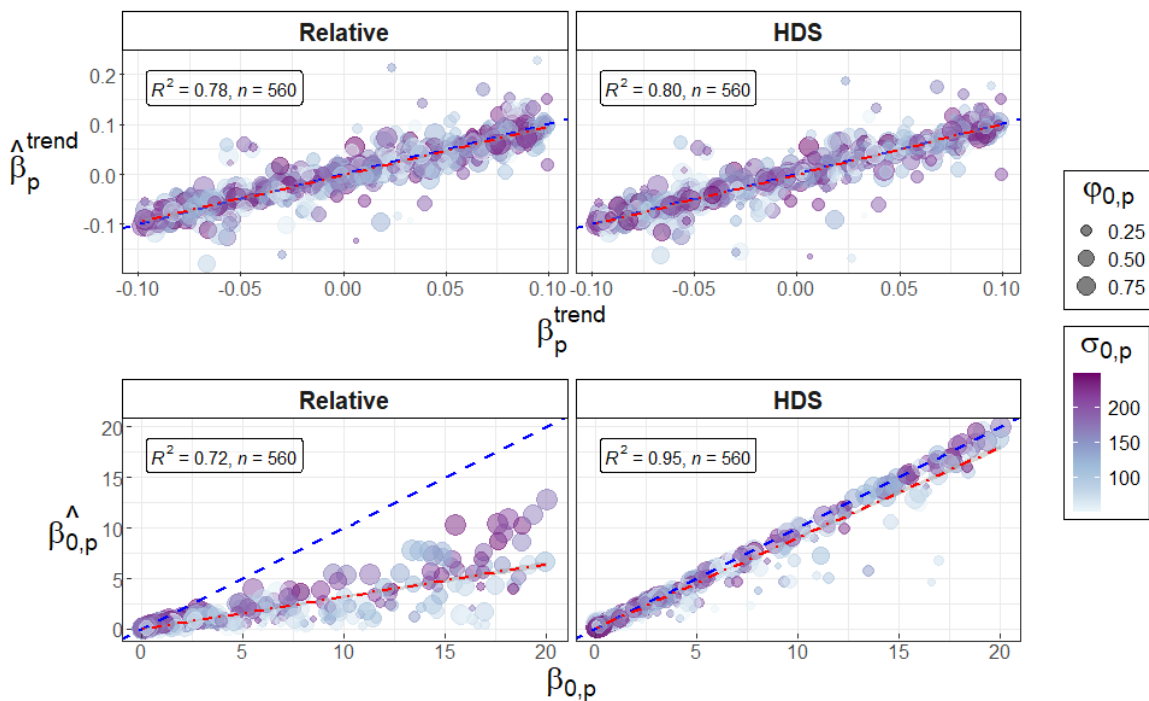
We assessed variation in CWMs, phylogenetic and taxonomic diversity between the two formulations using a linear model. We took account of the interaction between model formulation and the type of habitat surveyed (both categorical covariates) and year (continuous) as fixed effects, while also implementing random effects over observers. It corresponds to the following pseudo-R syntax :

$$Community\ index \sim Model * Habitats + Year + (1|Observer)$$

## Results

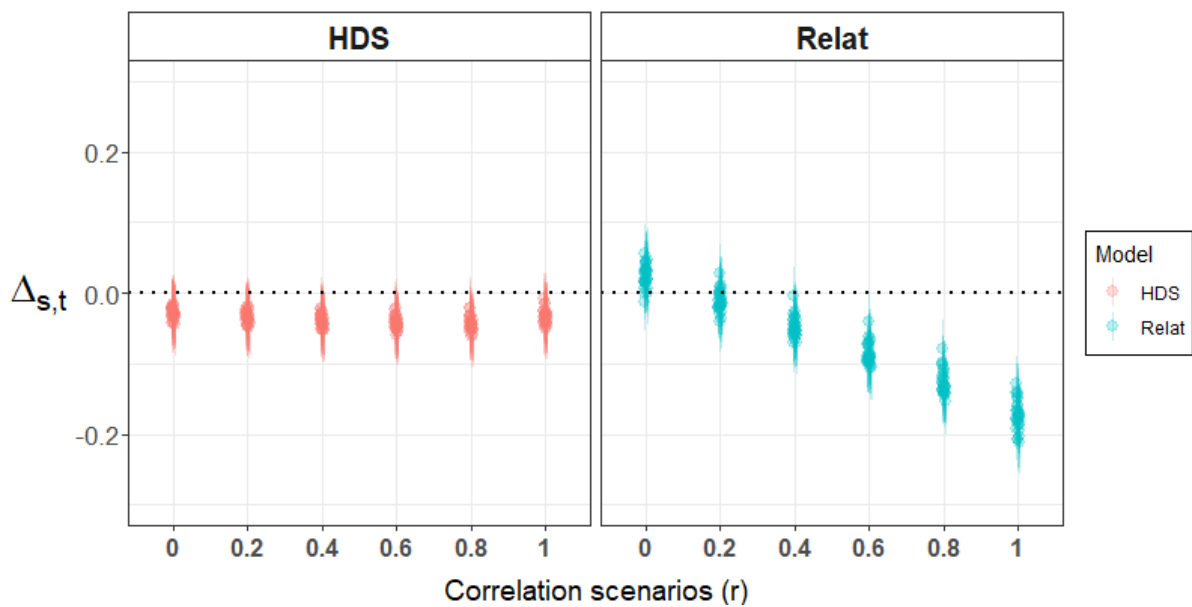
### Simulation study

Results from the simulation study (fig 4 - upper row) show that both  $\mathbf{M}_{\text{Relat}}$  and  $\mathbf{M}_{\text{HDS}}$  are capable of accurately estimating the simulated species trend ( $\beta_p^{\text{trend}}$ ). While  $\mathbf{M}_{\text{HDS}}$  correctly assessed simulated mean species abundance,  $\mathbf{M}_{\text{Relat}}$  systematically underestimated species abundances (Fig 4 - lower row), thus confirming **(H1)**, in a simulated environment. See appendix S4, for identifiability plots of parameters  $\beta$ , for both models, and  $\varphi_0, \sigma_0, \alpha$  and  $\gamma$ , for  $\mathbf{M}_{\text{HDS}}$ .



**Fig 4: Identifiability plot from results of the simulation study for ( $\beta_p^{\text{trend}}$  : upper row and  $\beta_{0,p}$  : lower row) over the natural scale (mean number of individuals). Blue line represents  $y=x$ , accurate estimation of simulated parameters. Red line depicts the linear trend from estimated and simulated parameters. Dot sizes correspond to species availability ( $\varphi_{0,p}$ : probability of being exposed to survey visits). Colour gradient represents species detectability ( $\sigma_{0,p}$ : the parameter determining the decreasing rate of the half-normal detection function), which estimates the probability of detecting an individual considered available given its distance of observation; given a set distance high values of ( $\sigma_{0,p}$ ) lead to higher probabilities of detections.**

When analysing the differences ( $\Delta_{s,t}$ ) between simulated and estimated CWM over different correlation ( $r$ ) between species detectability ( $\sigma_{0,p}$ ) and species traits ( $\theta_p$ ), we observed no discernible trend, with each estimated difference intersecting with 0, for  $\mathbf{M}_{\text{HDS}}$ . In contrast,  $\mathbf{M}_{\text{Relat}}$  depicted a negative trend between  $r$  and  $\Delta_{s,t}$ , with evident gap from 0 beyond  $r = 0.4$  (Fig 5). These results support **H2**, suggesting that communities composed of species whose traits affect their detectability could lead to biased community index.

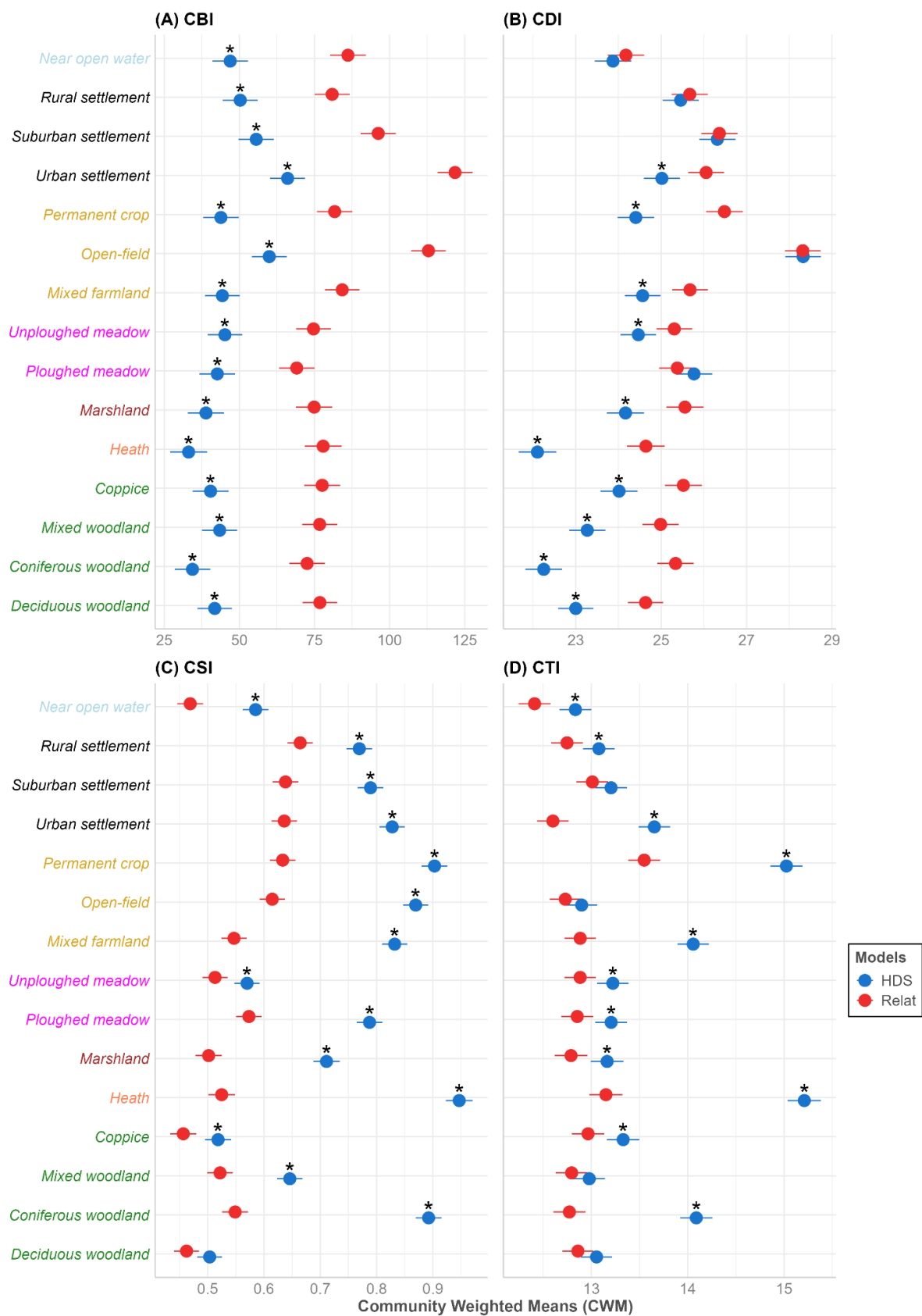


**Fig 5: Plot of the distribution of the difference ( $\Delta_{s,t}$ ) between simulated and estimated CWM over distinct correlation scenarios ( $r$ ) between species detectability and species traits. Each point and its associated uncertainty range derived from bootstrap resamples, with each resample corresponding to one interaction of the bootstrap analysis. Dotted line represent  $\Delta_{s,t} = 0$  corresponding to perfect estimation of  $CWM_{s,t}$ ; negative and positive values depict overestimation and underestimation of CWM.**

*Case study*

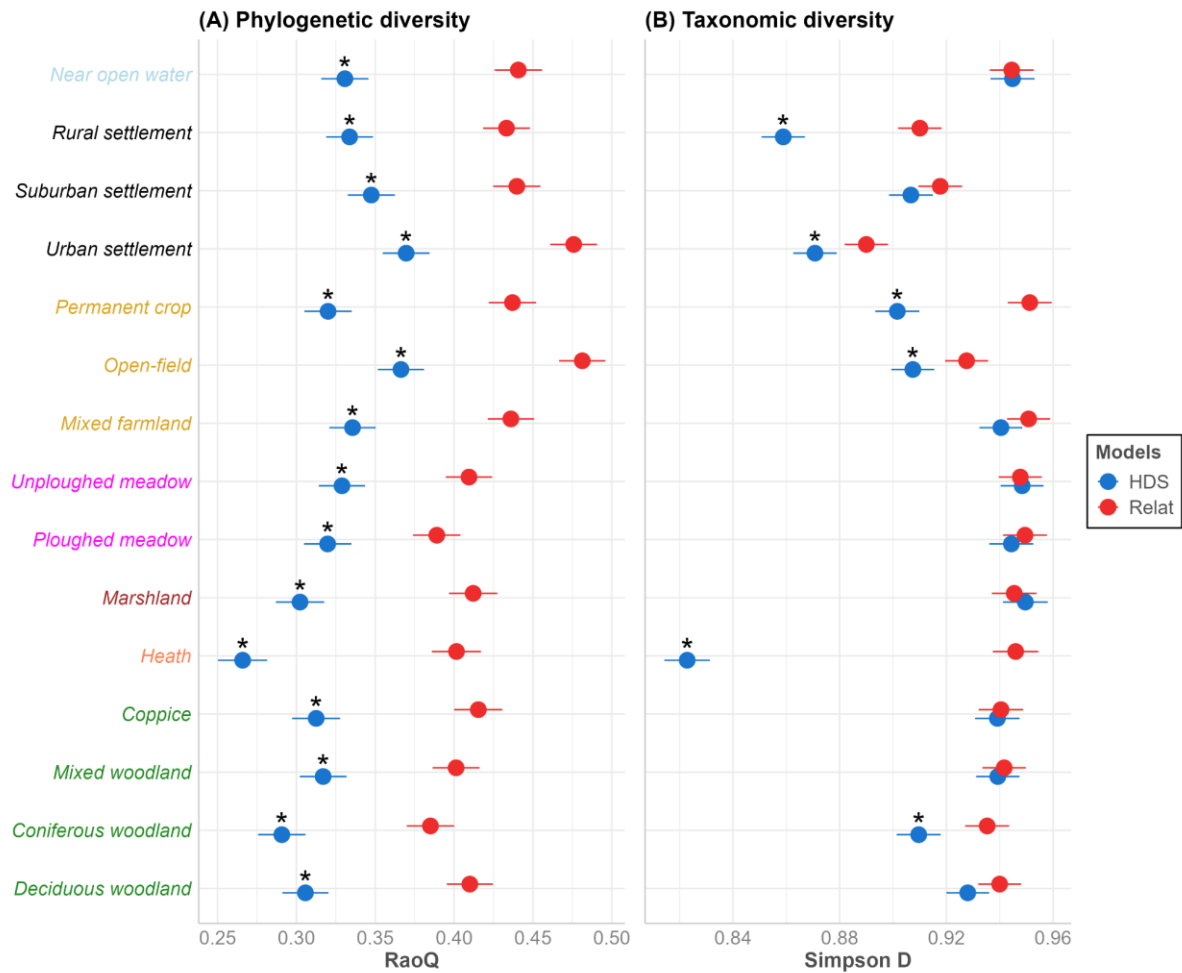
There are distinct response patterns from the marginal effect plots of Community-Weighted Body mass (CBI) produced from species weights ( $p_{s,t,p}$ ) estimated by  $\mathbf{M}_{\text{Relat}}$  and  $\mathbf{M}_{\text{HDS}}$  models (Fig 6A-D). Across all habitats, we had significantly lower values of CBI using species relative weights in communities ( $p_{s,t,p}$ ) estimated from  $\mathbf{M}_{\text{HDS}}$  than  $\mathbf{M}_{\text{Relat}}$  (Fig 6A). The same pattern appeared for Community Dispersal Index (CDI), based on species Hand-Wing index, with significantly lower estimates of CDI using species abundance from  $\mathbf{M}_{\text{HDS}}$ , despite habitats such as near open water, rural and suburban settlements, open-field and ploughed meadows (Fig 6B). In contrast, for community specialisation index (CSI) estimates from  $\mathbf{M}_{\text{HDS}}$  are significantly higher across all habitats (Fig 6C). Similarly, for the community thermal index (CTI),  $\mathbf{M}_{\text{HDS}}$  had significantly higher CTI estimates than  $\mathbf{M}_{\text{Relat}}$ , except in suburban settlement, open-field, mixed woodland and deciduous woodland habitats (Fig 6D).

For bird diversity metrics (Fig 7), the phylogenetic index (RaoQ) computed using species weights estimated from  $\mathbf{M}_{\text{HDS}}$  was significantly lower than  $\mathbf{M}_{\text{Relat}}$  ones (Fig 7A). This suggests that accounting for the observation process leads to estimations of greater phylogenetic dissimilarities (*i.e.* lower RaoQ). In contrast to taxonomic diversity assessed by the Simpson index (Fig 7B), where we had estimated equivalent diversity dissimilarity between  $\mathbf{M}_{\text{HDS}}$  and  $\mathbf{M}_{\text{Relat}}$ , apart from rural and urban settlement, permanent crop, open-field, heath and coniferous woodland habitat types with significantly lower estimated index. These results suggest that communities could be more dissimilar (*i.e.* lower Simpson D index) when accounting for the observation process.



**Fig 6: Marginal effect plots of estimated CMW, and their associated confidence intervals, over habitat types prospected by the FBBS schemes. In (A) Community Body mass Index (CBI); (B) Community Dispersal Index (CDI) using species Hand-Wing Index (HWI), (C) Community Specialization Index (CSI) and (D) Community**

**Thermal Index (CTI).** Red and blue colours correspond to results from  $M_{Relat}$  and  $M_{HDS}$  model formulation respectively. No overlap of marginal means confidence intervals ( $\alpha = 0.05$ ) are depicted by black asterisks.



**Fig 7: Marginal effect plot of diversity indices, over habitat types, using species weight ( $p_{s,t,p}$ ) estimated from either  $M_{Relat}$  (red) and  $M_{HDS}$  (blue) models. No overlap of marginal means confidence intervals ( $\alpha = 0.05$ ) are depicted by black asterisks.**

## Discussion

Our simulation results confirmed our two main hypotheses, that omission of the observation process for the estimation of species abundance could lead to substantial biased estimates of community weighted means (CWM) indices (**H1**). They also highlight substantial overestimation of community indices when communities consist of species whose traits are strongly ( $r \geq 0.5$ ) associated with their detectability (**H2**). Results from our case study refine these simulations, by showing significant differences with and without

accounting for the observation process for the estimation of CWMs related to species traits and functions affected by detectability (**H2**), as well as for diversity indices (**H3**). Our results suggest the potential need to reconsider our understanding of the bird assemblage composition across multiple habitats in France and demonstrates that accounting for the observation process is crucial for community studies even for the analysis of standardised data collection.

Our results highlight major differences, across multiple habitats, between CWMs estimated by  $M_{HDS}$  and  $M_{Relat}$ , especially for traits based on species morphometry, *e.g.* body mass and Hand-Wing Index (HWI). Given the positive relation between species detectability and species body mass (Jachmann, 2002; Johnston et al., 2014), lower CWM estimated from  $M_{HDS}$  could suggest that a non negligible quantity of small birds are unaccounted for when omitting the observation process. We see a similar response pattern for HWI, while being negatively correlated to species body mass (Fig 4A; Sheard et al., 2020). As the HWI corresponds to a proxy of flight efficiency and species dispersal capabilities (Sheard et al., 2020), this could imply potential biased estimates of communities functions, here dispersal capabilities, when the observation process is unaccounted for (Palacio et al., 2020).

For this study, in order to conform to the distance sampling assumption of measurement of exact observation distance (Buckland 1993), we decided to remove observation of individuals detected in flight to filter out potential misinformed observation distance. This approach required us to exclude four species with high HWI, *i.e.*, two migratory urban species, the Common swift (*Apus apus*) and Barn swallow (*Hirundo rustica*) and two birds of prey, the Common kestrel (*Falco tinnunculus*) and Common buzzard (*Buteo buteo*).

In the case of species specialisation, higher CSI estimates from **M<sub>HDS</sub>**, could suggest that specialist species, *i.e.* species exploiting a smaller range of habitat, are predominantly missed during surveys (Mackenzie et al., 2005). These biased estimates of specialist species undermine their contribution to the community indices, leading to misrepresentation of communities (Iknayan et al., 2014).

For the Community Thermal Index (CTI), we considered the species thermal index (STI) to have minimal correlation to detectability, *i.e.*  $r$  near zero in our simulation cases. We based this assumption on the primary aim of this index intending to represent the average temperature experienced across species distribution range (Devictor et al., 2008a). This index is focused on studying species distribution range shifts and assemblage composition driven by climate change (Barnagaud et al., 2012; Devictor et al., 2008a; Gaüzère et al., 2020), rather than behavioural responses to short-term and extreme temperature anomaly (Cohen et al., 2020) which could induce variation in detection probabilities (Alldredge et al., 2007). We did not find studies linking directly STI to species detection probabilities; however, it was shown that temperature variation caused by climate change could affect breeding dates (Lehikoinen and Sparks, 2010) potentially resulting in variation of species detectability across the breeding season (Emlen, 1977; Garson and Hunter, 1979; Strebel et al., 2014).

For phylogenetic and taxonomic diversity, our results corroborate with previous studies (Jarzyna and Jetz, 2016; Richter et al., 2021; Roth et al., 2018; Si et al., 2018) where unaccounted observation process lead to biased diversity indices, due to the phylogenetic distribution of species traits affecting species detectability (Maurer, 1998; Sheard et al., 2020; Sóllymos et al., 2018).

Despite development of practices towards collection of adequate data allowing modelling of the observation process (Rosenstock et al., 2002; Yoccoz et al., 2001) their

integration into ecological studies is still relatively modest. (Kellner and Swihart, 2014). Here, it can be transcribed through prior instances of studies using the same schemes as this study. We relied on the FBBS (French Breeding Bird Survey) which was profusely used in previous studies over bird communities that either (i) did not account for the detection process, leading to potential biased communities indices estimation (Dross et al., 2018; Filippi-Codaccioni et al., 2010; Gaüzère and Devictor, 2021; Jiguet et al., 2010; Lorel et al., 2021; Princé et al., 2012; Vallé et al., 2023), (ii) did not account for it, while discussing resulting potential biases (Barnagaud et al., 2012; Godet et al., 2016; Monnet et al., 2014; Pellissier et al., 2013) and (iii) studies integrating it (Devictor et al., 2007; Doxa et al., 2010; Giraud et al., 2016; Jiguet et al., 2005; Rigal et al., 2022). Given our results, this could suggest that studies corresponding to (i) and (ii) may yield biased estimates of community indices, which might have lead to a potential misrepresentation of changes in community composition (Iknayan et al., 2014; Palacio et al., 2020; Tingley and Beissinger, 2013). It shows that there is a persisting reliance upon the standardisation of the sampling scheme expecting that it should capture detection error originated from the observation process (Bonar et al., 2010). Though, regardless of tedious data collection efforts, unaccounted variation in species detection might still arise (Hochachka and Fiedler, 2008; Schmidt, 2005).

Such unaccounted variations stemming from multiple sources, such as the gain of observers expertise across the years (Riddle et al., 2010), first-year survey effect (Kendall et al., 1996), species morphometric (Jachmann, 2002; Johnston et al., 2014), between-year variation induced by phenological shifts (Devictor et al., 2012) can be accounted for by Hierarchical Models (Bellier et al., 2023; Dorazio et al., 2006; Tingley et al., 2020). As CWMs and most of diversity indices are based on species abundances (Gaüzère et al., 2019; Maurer and McGill, 2010), the violation of the proportionality assumption between detected and actual number of individuals (Thompson, 2002) of inferred abundance propagates biases

resulting in community indices biased towards more easily detectable species (Palacio et al., 2020). HM flexibility could also improve community assessment through data integration, *i.e.* inclusion of additional data collected from diverse data sources (Fletcher et al., 2019; Pacifici et al., 2017; Zipkin et al., 2021) permitting (i) improvement of habitat- and detection-related estimates and (ii) allowing depiction of population structure using Integrated Population Models (IPM ; Baillie, 1990; Schaub and Abadi, 2011) approaches. Here, we relied on point count distance data without sexing of detected individuals. This results in inferred abundance under the assumption that male individuals are mostly detected during survey (Catchpole and Slater, 2008; Odom et al., 2014; Riebel et al., 2019). Additional information of species population structures, *e.g.* sex ratio or age-structure, could help refine estimation of CWM, for instance using differentiated body mass between male and female (Caron and Pie, 2024; Székely et al., 2007) as well as additional morphometric data collected from birds traps and nest surveys schemes (Noon et al., 1992; Robinson et al., 2009).

Our results stress the need to account for the observation process in ecological studies of bird communities. Species with diverse life-history traits result in uneven numbers detected during sampling occasion among trait values, leading to biased estimates of species abundances and their relative weights in communities. Our simulation study showed a negative trend between community indices biases and categories of species traits, according to their respective detectability. These results are supported by our case study based on a multi-year (2015-2023) field survey over metropolitan France suggesting a potential misrepresentation of taxonomic, phylogenetic and functional community composition, despite a data collection relying on a standardised scheme.

## V.C Supplementary information

*S1: MCMC parameters and priors*

### ***Simulation study***

#### MCMC parameters

Iteration: 6,000

Burn-in: 3,000

Thin: 3

Chains: 4

#### Range for species parameter generation

$$\beta_{0,p} \sim U(-3,3)$$

$$\beta_p \sim U(-2,2)$$

$$\beta_p^{trend} \sim U(-0.1,0.1)$$

$$\bar{\varphi}_{0,p} \sim U(0.05,1) \text{ and } \varphi_{0,p} = \text{logit}(\bar{\varphi}_{0,p})$$

$$\gamma_p \sim U(-2,2)$$

$$\bar{\sigma}_{0,p} \sim U(50,250) \text{ and } \sigma_{0,p} = \log(\bar{\sigma}_{0,p})$$

$$\alpha_p \sim U(-2,2)$$

#### MCMC Prior

$$\beta_{0,p} \sim \text{Normal}(0,10)$$

$$\beta_p \sim \text{Normal}(0,10)$$

$$\beta_p^{\text{trend}} \sim U(-1,1)$$

$$\bar{\varphi}_{0,p} \sim U(0,1) \text{ and } \varphi_{0,p} = \text{logit}(\bar{\varphi}_{0,p})$$

$$\bar{\sigma}_{0,p} \sim U(0,500) \text{ and } \sigma_{0,p} = \log(\bar{\sigma}_{0,p})$$

$$\gamma_p \sim \text{Normal}(0,10)$$

$$\alpha_p \sim \text{Normal}(0,10)$$

### ***Case study***

#### MCMC parameters

Iteration: 20 000

Burn-in: 10 000

Thin: 40

Chains: 4

#### Prior

$$\beta_{0,p} \sim \text{Normal}(0,10)$$

$$\beta_p \sim \text{Normal}(0,10)$$

$$\beta_p^{\text{trend}} \sim U(-1,1)$$

$$\sum_{h=1}^H \beta_{h,p}^{\text{cat}} \sim \text{Normal}(0,10)$$

$$\bar{\sigma}_{0,p} \sim U(0,500) \text{ and } \sigma_{0,p} = \log(\bar{\sigma}_{0,p})$$

$$\varphi_{0,p} \sim U(-10,10)$$

$$\sum_{a=1}^4 \alpha_{a;p} \sim \text{Normal}(0,10)$$

$$\varepsilon_s^{abund} \sim \text{Normal}(\varepsilon_g, \tau_g)$$

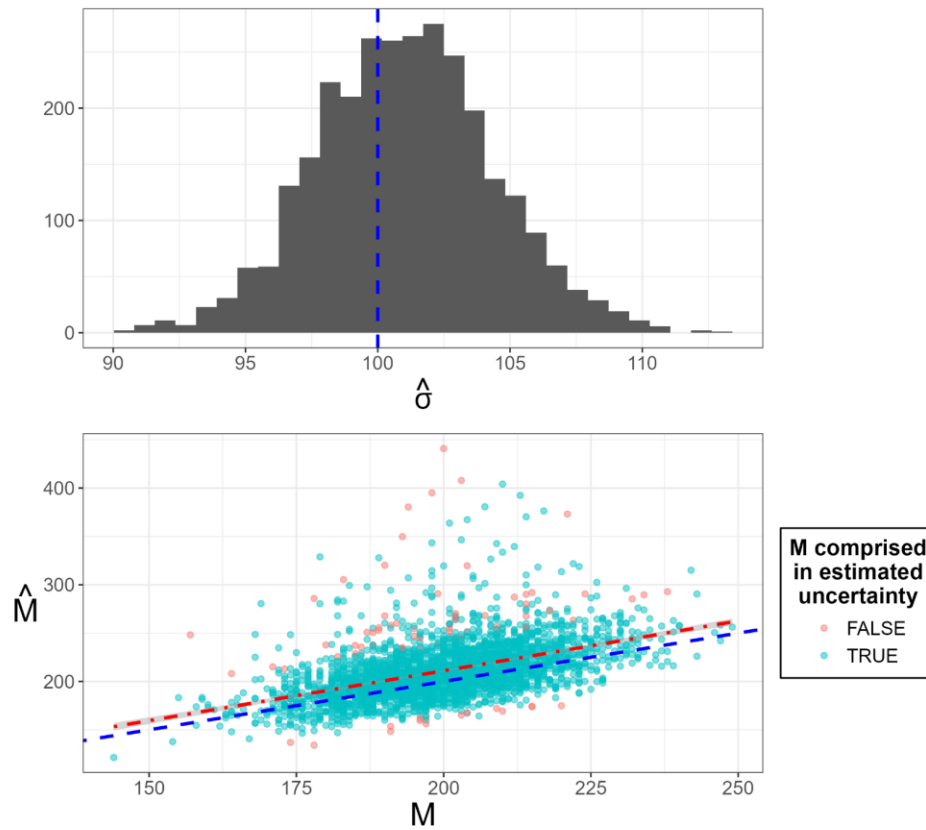
$$\sigma_g \sim \text{Gamma}(1,1)$$

$$\varepsilon_g \sim \text{Normal}(0, \tau_g)$$

*S2: Bias assessment for coarse distance classes intervals*

To account for the scheme unequal distance bins collection (A = 0-25; B = 25-50; C = 50-100; D = 100-200; E = 200-200+), we needed to computed detection function cell probabilities over finer and equal distance bins. For the case study analysis, we chose finer distance bins of 25m, that were summed to abide to real data collection (A = 0-25; B = 25-50; C = 50-75 and 75-100; D = 100-125, 125-150, 150-175 and 175-200; E = 200-225, 225-250, 250-275, 275-300, 325-350, 350-375 and 375-400). As the right-side of the last distance bins (E) is left open, we decided to limit it to 400m, based on prior estimates of maximal distance of observation (after truncation of 5% of extreme values) collected from independent data (Nabias et al., 2024).

Here, we estimated the detection function parameter ( $\sigma_0$ ) using ad-hoc finer distance bins generated during detection function modelling. We simulated 3000 datasets comprising point count distance data collected over 100 sites surveyed 9 times in one year. Across simulated datasets, we used constant parameters of species mean abundances ( $\beta_0 = 2$ ), species availability, *i.e.* species probability of being exposed to visit occasion, ( $\varphi_0 = 1$ ; here species are always considered exposed to observers visits) and species detectability ( $\sigma_0 = 100$ ), linked to species probabilities of being detected by observers while being considered exposed. For this simulation study, we used 20.000 iterations, with 10.000 burn-in over 4 chains for the Monte Carlo Markov Chain (MCMC) calibrated using JAGS 3.4.1 (Plummer, 2003).



**Fig S2.1: Identifiability plot for species detectability parameter ( $\sigma_0$ ) and estimated number of individuals ( $M$ ). For  $\sigma_0$ , we depicted the histogram of estimated ( $\hat{\sigma}_0$ ) values, as  $\sigma_0$  was held constant at 100, the vertical blue dashed-line, across all simulations. For  $M$ , the blue dashed-line represents  $y=x$ , *i.e.* accurate estimation of simulated parameters. Red line depicts the linear trend from estimated ( $\hat{M}$ ) and simulated  $M$  values.**

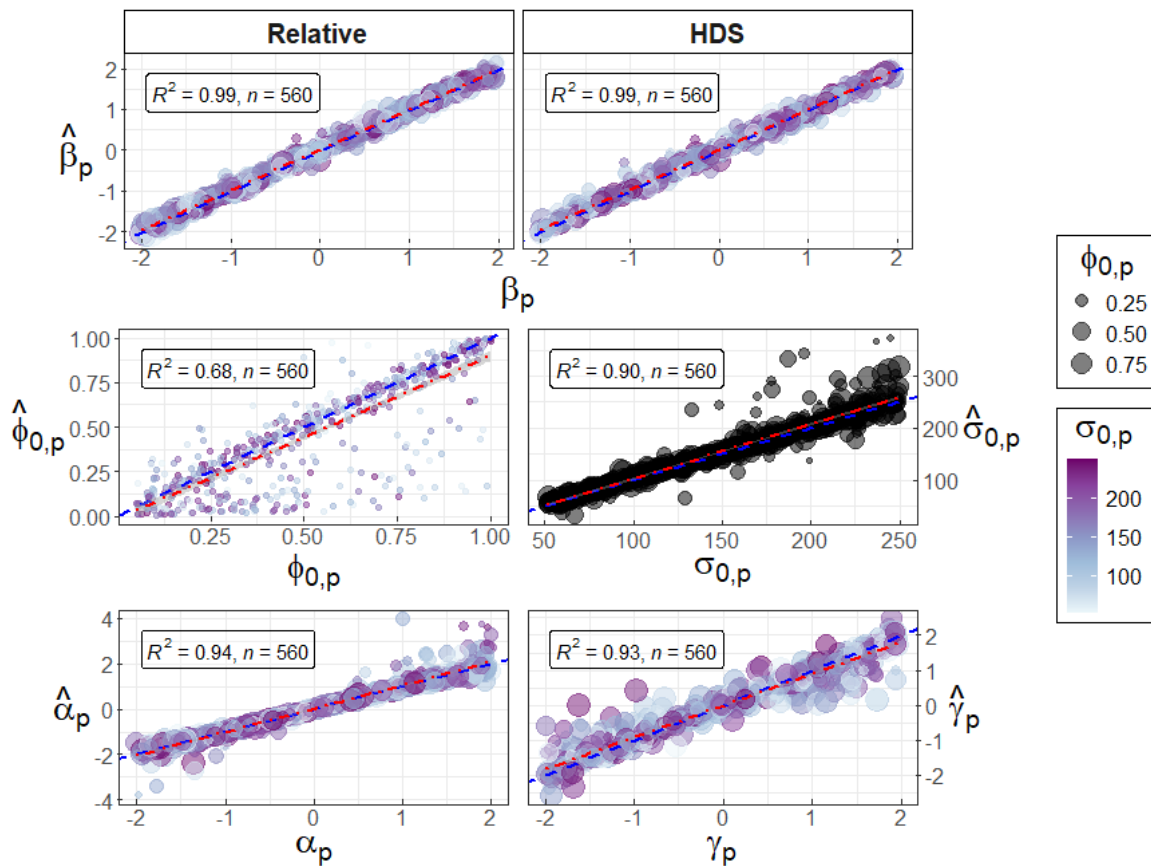
S3: Bird traits data

**Table S2.1: Species traits data used for studied communities indices. Species Thermal Index (STI ; Devictor et al., 2008) and Species Specialisation Index (SSI ; Julliard et al., 2006) are derived from data collected by the French Breeding Bird Survey (FBBS). We also used species body mass and Hand-Wing Index (HWI) from AVONET database (Tobias et al., 2022).**

English	Scientific names	Species Thermal Index (STI)	Species Specialisation Index (SSI)	Body mass	Hand-Wing Index (HWI)
Hedge Accentor	<i>Prunella modularis</i>	10,99	0,42	20,2	21,2
Eurasian Skylark	<i>Alauda arvensis</i>	12,51	1,19	37,3	33,6
Wood Lark	<i>Lullula arborea</i>	14,51	0,91	26,9	27,7
Yellow Wagtail	<i>Motacilla flava</i>	12,62	1,86	17,7	31,1
Yellowhammer	<i>Emberiza citrinella</i>	11,75	0,68	29,7	25,7
Common Buzzard	<i>Buteo buteo</i>	13,02	0,57	759,1	38,3
Common Quail	<i>Coturnix coturnix</i>	14,69	1,40	96,3	41,0
European Goldfinch	<i>Carduelis carduelis</i>	14,69	0,75	16,0	33,4
Common Cuckoo	<i>Cuculus canorus</i>	12,71	0,39	111,4	43,9
Common Kestrel	<i>Falco tinnunculus</i>	13,76	0,63	183,2	51,0
Blackcap	<i>Sylvia atricapilla</i>	13,19	0,29	16,7	27,0
Common Whitethroat	<i>Sylvia communis</i>	13,39	0,62	15,1	24,2
Sardinian Warbler	<i>Sylvia melanocephala</i>	18,33	1,62	11,7	18,0
Eurasian Jay	<i>Garrulus glandarius</i>	12,93	0,45	159,5	19,4
Eurasian Treecreeper	<i>Certhia familiaris</i>	11,77	1,96	9,0	19,2
Short-toed Treecreeper	<i>Certhia brachydactyla</i>	15,36	0,61	8,2	17,7
Mistle Thrush	<i>Turdus viscivorus</i>	12,44	0,53	117,4	35,1
Song Thrush	<i>Turdus philomelos</i>	11,60	0,45	67,7	30,7
Barn Swallow	<i>Hirundo rustica</i>	13,72	0,76	17,9	52,6
Melodious Warbler	<i>Hippolais polyglotta</i>	16,20	0,54	11,0	21,2
Eurasian Linnet	<i>Carduelis cannabina</i>	14,21	0,68	19,5	34,4
Common Swift	<i>Apus apus</i>	13,23	1,27	37,6	71,6
Eurasian Blackbird	<i>Turdus merula</i>	13,71	0,24	102,7	23,1
Blue Tit	<i>Parus caeruleus</i>	13,82	0,36	11,1	21,0
Great Tit	<i>Parus major</i>	12,89	0,28	16,3	18,7
Crested Tit	<i>Parus cristatus</i>	12,19	1,54	11,0	19,8
Coal Tit	<i>Parus ater</i>	12,33	1,57	9,2	21,9
Marsh Tit	<i>Parus palustris</i>	13,36	1,08	11,1	17,9
House Sparrow	<i>Passer domesticus</i>	13,89	1,28	26,5	27,3
Eurasian Tree Sparrow	<i>Passer montanus</i>	12,82	1,07	21,4	24,6
Grey Partridge	<i>Perdix perdix</i>	13,39	2,20	405,3	33,8
Red-legged Partridge	<i>Alectoris rufa</i>	15,24	1,11	527,9	25,8
Great Spotted Woodpecker	<i>Dendrocopos major</i>	12,19	0,64	74,9	29,3
Middle Spotted Woodpecker	<i>Dendrocopos medius</i>	14,63	2,22	59,0	28,1
Black Woodpecker	<i>Dryocopus martius</i>	11,39	1,07	321,0	26,4

Eurasian Green Woodpecker	<i>Picus viridis</i>	13,78	0,34	176,0	24,7
Black-billed Magpie	<i>Pica pica</i>	12,68	0,73	217,5	22,9
Common Wood-pigeon	<i>Columba palumbus</i>	12,99	0,34	490,0	41,2
Eurasian Chaffinch	<i>Fringilla coelebs</i>	13,07	0,29	23,8	27,6
Meadow Pipit	<i>Anthus pratensis</i>	9,85	1,38	18,4	24,3
Wood Warbler	<i>Phylloscopus sibilatrix</i>	12,47	1,82	9,2	30,9
Common Chiffchaff	<i>Phylloscopus collybita</i>	11,58	0,47	8,3	18,4
Firecrest	<i>Regulus ignicapilla</i>	14,55	1,14	5,6	22,0
Goldcrest	<i>Regulus regulus</i>	11,19	1,67	5,5	22,0
Common Nightingale	<i>Luscinia megarhynchos</i>	16,01	0,48	19,6	27,6
European Robin	<i>Erithacus rubecula</i>	12,24	0,52	17,7	19,9
Common Redstart	<i>Phoenicurus phoenicurus</i>	11,77	0,76	14,6	25,8
Black Redstart	<i>Phoenicurus ochruros</i>	14,46	1,06	16,5	23,1
European Serin	<i>Serinus serinus</i>	15,55	0,90	11,2	29,3
Wood Nuthatch	<i>Sitta europaea</i>	13,02	0,96	20,4	23,8
Whinchat	<i>Saxicola rubetra</i>	11,66	1,60	16,6	26,7
Common Stonechat	<i>Saxicola torquatus</i>	15,21	0,80	14,1	20,8
Eurasian Collared-dove	<i>Streptopelia decaocto</i>	14,66	1,04	149,0	35,4
Northern Wheatear	<i>Oenanthe oenanthe</i>	11,59	1,89	25,4	32,1
Winter Wren	<i>Troglodytes troglodytes</i>	12,91	0,43	9,7	15,7

S4: Additional identifiability plots of species parameters



**Fig S4.1: Identifiability plot of species parameters.** While identifiability plots for  $(\beta_p; \text{upper row})$  are depicted for both model formulation,  $(\varphi_{0,p}; \text{middle-left}, \sigma_{0,p}; \text{middle-right}, \alpha_p; \text{bottom-left and } \gamma_p; \text{bottom-right})$  are estimated only for  $M_{\text{HDS}}$ . Blue line represents  $y=x$ , accurate estimation of simulated parameters. Red line depicts the linear trend from estimated and simulated parameters. Dot sizes correspond to species availability  $(\varphi_{0,p}$ : probability of being exposed to survey visits). Colour gradient represents species detectability  $(\sigma_{0,p}$ : the parameter determining the decreasing rate of the half-normal detection function). It estimates the probability of detecting an individual considered available given its distance of observation; given a set distance, higher values of  $\sigma_{0,p}$  lead to higher probabilities of detections.

## V.D References

- Ali, A., 2023. Biodiversity–ecosystem functioning research: Brief history, major trends and perspectives. *Biol. Conserv.* 285, 110210. <https://doi.org/10.1016/j.biocon.2023.110210>
- Allredge, M.W., Simons, T.R., Pollock, K.H., 2007. FACTORS AFFECTING AURAL DETECTIONS OF SONGBIRDS. *Ecol. Appl.* 17, 948–955. <https://doi.org/10.1890/06-0685>
- Arango, A., Pinto-Ledezma, J., Rojas-Soto, O., Lindsay, A.M., Mendenhall, C.D., Villalobos, F., 2022. Hand-Wing Index as a surrogate for dispersal ability: the case of the Emberizoidea (Aves: Passeriformes) radiation. *Biol. J. Linn. Soc.* 137, 137–144. <https://doi.org/10.1093/biolinnean/blac071>
- Baillie, S.R., 1990. Integrated population monitoring of breeding birds in Britain and Ireland. *Ibis* 132, 151–166. <https://doi.org/10.1111/j.1474-919X.1990.tb01035.x>
- Barnagaud, J.-Y., Devictor, V., Jiguet, F., Barbet-Massin, M., Viol, I.L., Archaux, F., 2012. Relating Habitat and Climatic Niches in Birds. *PLOS ONE* 7, e32819. <https://doi.org/10.1371/journal.pone.0032819>
- Begon, M., Townsend, C.R., Harper, J.L., 2006. *Ecology: from individuals to ecosystems*, 4th ed. ed. Blackwell Pub, Malden, MA.
- Bellier, E., Rectenwald, J., Sisson, C., Terhune, T.M., Martin, J.A., 2023. Estimating population dynamics trajectories of raptors from a multi-species hierarchical distance sampling model. *Ecol. Inform.* 75, 102024. <https://doi.org/10.1016/j.ecoinf.2023.102024>
- Blowes, S.A., McGill, B., Brambilla, V., Chow, C.F.Y., Engel, T., Fontrodona-Eslava, A., Martins, I.S., McGlenn, D., Moyes, F., Sagouis, A., Shimadzu, H., van Klink, R., Xu, W.-B., Gotelli, N.J., Magurran, A., Dornelas, M., Chase, J.M., 2024. Synthesis reveals approximately balanced biotic differentiation and homogenization. *Sci. Adv.* 10, eadj9395. <https://doi.org/10.1126/sciadv.adj9395>
- Bonar, S.A., Fehmi, J.S., Mercado-Silva, N., 2010. An overview of sampling issues in species diversity and abundance surveys, in: Magurran, A.E., McGill, B.J. (Eds.), *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford University Press, Oxford, New York, pp. 11–24.
- Buckland, S.T., Rexstad, E.A., Marques, T.A., Oedekoven, C.S., 2015. *Distance Sampling: Methods and Applications*, *Methods in Statistical Ecology*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-19219-2>
- Burleigh, J.G., Kimball, R.T., Braun, E.L., 2015. Building the avian tree of life using a large-scale, sparse supermatrix. *Mol. Phylogenet. Evol.* 84, 53–63. <https://doi.org/10.1016/j.ympev.2014.12.003>
- Cadotte, M.W., Carscadden, K., Mirotchnick, N., 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. *J. Appl. Ecol.* 48, 1079–1087. <https://doi.org/10.1111/j.1365-2664.2011.02048.x>
- Caron, F.S., Pie, M.R., 2024. The macroevolution of sexual size dimorphism in birds. *Biol. J. Linn. Soc.* blad168. <https://doi.org/10.1093/biolinnean/blad168>
- Catchpole, C.K., Slater, P.J.B., 2008. *Recognition and territorial defence*, in: *Bird Song: Biological Themes and Variations*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511754791>

- Chen, G., Kéry, M., Plattner, M., Ma, K., Gardner, B., 2013. Imperfect detection is the rule rather than the exception in plant distribution studies. *J. Ecol.* 101, 183–191. <https://doi.org/10.1111/1365-2745.12021>
- Claramunt, S., Derryberry, E.P., J. V. Remsen, J., Brumfield, R.T., 2012. High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proc. R. Soc. B Biol. Sci.* <https://doi.org/10.1098/rspb.2011.1922>
- Clavel, J., Julliard, R., Devictor, V., 2011. Worldwide decline of specialist species: toward a global functional homogenization? *Front. Ecol. Environ.* 9, 222–228. <https://doi.org/10.1890/080216>
- Cohen, J.M., Fink, D., Zuckerberg, B., 2020. Avian responses to extreme weather across functional traits and temporal scales. *Glob. Change Biol.* 26, 4240–4250. <https://doi.org/10.1111/gcb.15133>
- Devictor, V., Julliard, R., Couvet, D., Jiguet, F., 2008a. Birds are tracking climate warming, but not fast enough. *Proc. R. Soc. B Biol. Sci.* 275, 2743–2748. <https://doi.org/10.1098/rspb.2008.0878>
- Devictor, V., Julliard, R., Couvet, D., Lee, A., Jiguet, F., 2007. Functional Homogenization Effect of Urbanization on Bird Communities. *Conserv. Biol.* 21, 741–751. <https://doi.org/10.1111/j.1523-1739.2007.00671.x>
- Devictor, V., Julliard, R., Jiguet, F., 2008b. Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos* 117, 507–514. <https://doi.org/10.1111/j.0030-1299.2008.16215.x>
- Devictor, V., van Swaay, C., Brereton, T., Brotons, L., Chamberlain, D., Heliölä, J., Herrando, S., Julliard, R., Kuussaari, M., Lindström, Å., Reif, J., Roy, D.B., Schweiger, O., Settele, J., Stefanescu, C., Van Strien, A., Van Turnhout, C., Vermouzek, Z., WallisDeVries, M., Wynhoff, I., Jiguet, F., 2012. Differences in the climatic debts of birds and butterflies at a continental scale. *Nat. Clim. Change* 2, 121–124. <https://doi.org/10.1038/nclimate1347>
- Dorazio, R.M., Royle, J.A., Söderström, B., Glimskär, A., 2006. Estimating Species Richness and Accumulation by Modeling Species Occurrence and Detectability. *Ecology* 87, 842–854. [https://doi.org/10.1890/0012-9658\(2006\)87\[842:ESRAAB\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[842:ESRAAB]2.0.CO;2)
- Doxa, A., Bas, Y., Paracchini, M.L., Pointereau, P., Terres, J.-M., Jiguet, F., 2010. Low-intensity agriculture increases farmland bird abundances in France. *J. Appl. Ecol.* 47, 1348–1356. <https://doi.org/10.1111/j.1365-2664.2010.01869.x>
- Dross, C., Princé, K., Jiguet, F., Tichit, M., 2018. Contrasting bird communities along production gradients of crops and livestock in French farmlands. *Agric. Ecosyst. Environ.* 253, 55–61. <https://doi.org/10.1016/j.agee.2017.10.025>
- Emlen, J.T., 1977. ESTIMATING BREEDING SEASON BIRD DENSITIES FROM TRANSECT COUNTS. *Auk* 94, 455–468.
- Filippi-Codaccioni, O., Devictor, V., Bas, Y., Julliard, R., 2010. Toward more concern for specialisation and less for species diversity in conserving farmland biodiversity. *Biol. Conserv.* 143, 1493–1500. <https://doi.org/10.1016/j.biocon.2010.03.031>
- 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* 100, e02710. <https://doi.org/10.1002/ecy.2710>

- Garson, P.J., Hunter, M.L., 1979. Effects of temperature and time of year on the singing behavior of wrens *Troglodytes troglodytes* and great tits *Parus major*. *Ibis* 121, 481–487. <https://doi.org/10.1111/j.1474-919X.1979.tb06688.x>
- Gaüzère, P., Barbaro, L., Calatayud, F., Princé, K., Devictor, V., Raison, L., Sirami, C., Balent, G., 2020. Long-term effects of combined land-use and climate changes on local bird communities in mosaic agricultural landscapes. *Agric. Ecosyst. Environ.* 289, 106722. <https://doi.org/10.1016/j.agee.2019.106722>
- Gaüzère, P., Devictor, V., 2021. Mismatches between birds' spatial and temporal dynamics reflect their delayed response to global changes. *Oikos* 130, 1284–1296. <https://doi.org/10.1111/oik.08289>
- Gaüzère, P., Doucier, G., Devictor, V., Kéfi, S., 2019. A framework for estimating species-specific contributions to community indicators. *Ecol. Indic.* 99, 74–82. <https://doi.org/10.1016/j.ecolind.2018.11.069>
- Giraud, C., Calenge, C., Coron, C., Julliard, R., 2016. Capitalizing on Opportunistic Data for Monitoring Relative Abundances of Species. *Biometrics* 72, 649–658. <https://doi.org/10.1111/biom.12431>
- Godet, L., Devictor, V., Burel, F., Robin, J.-G., Ménanteau, L., Fournier, J., 2016. Extreme Landscapes Decrease Taxonomic and Functional Bird Diversity but Promote the Presence of Rare Species. *Acta Ornithol.* 51, 23–38. <https://doi.org/10.3161/00016454AO2016.51.1.003>
- Hochachka, W.M., Fiedler, W., 2008. Trends in trappability and stop-over duration can confound interpretations of population trajectories from long-term migration ringing studies. *J. Ornithol.* 149, 375–391. <https://doi.org/10.1007/s10336-008-0282-1>
- Iknayan, K.J., Tingley, M.W., Furnas, B.J., Beissinger, S.R., 2014. Detecting diversity: emerging methods to estimate species diversity. *Trends Ecol. Evol.* 29, 97–106. <https://doi.org/10.1016/j.tree.2013.10.012>
- Jachmann, H., 2002. Comparison of aerial counts with ground counts for large African herbivores. *J. Appl. Ecol.* 39, 841–852. <https://doi.org/10.1046/j.1365-2664.2002.00752.x>
- Jarzyna, M.A., Jetz, W., 2016. Detecting the Multiple Facets of Biodiversity. *Trends Ecol. Evol.* 31, 527–538. <https://doi.org/10.1016/j.tree.2016.04.002>
- Jiguet, F., Devictor, V., Julliard, R., Couvet, D., 2012. French citizens monitoring ordinary birds provide tools for conservation and ecological sciences. *Acta Oecologica, Sino-French Cooperation on Biodiversity and Natural Resource Management* 44, 58–66. <https://doi.org/10.1016/j.actao.2011.05.003>
- Jiguet, F., Devictor, V., Ottvall, R., Turnhout, C.V., Jeugd, H.V. der, Lindström, Å., 2010. Bird population trends are linearly affected by climate change along species thermal ranges. *Proc. R. Soc. B Biol. Sci.* <https://doi.org/10.1098/rspb.2010.0796>
- Jiguet, F., Renault, O., Petiau, A., 2005. Estimating species richness with capture–recapture models: choice of model when sampling in heterogeneous conditions. *Bird Study* 52, 180–187. <https://doi.org/10.1080/00063650509461389>
- Johnston, A., Newson, S.E., Risely, K., Musgrove, A.J., Massimino, D., Baillie, S.R., Pearce-Higgins, J.W., 2014. Species traits explain variation in detectability of UK birds. *Bird Study* 61, 340–350. <https://doi.org/10.1080/00063657.2014.941787>

- Julliard, R., Clavel, J., Devictor, V., Jiguet, F., Couvet, D., 2006. Spatial segregation of specialists and generalists in bird communities. *Ecol. Lett.* 9, 1237–1244. <https://doi.org/10.1111/j.1461-0248.2006.00977.x>
- Kellner, K., Meredith, M., 2021. jagsUI: A Wrapper Around “rjags” to Streamline “JAGS” Analyses.
- Kellner, K.F., Swihart, R.K., 2014. Accounting for Imperfect Detection in Ecology: A Quantitative Review. *PLOS ONE* 9, e111436. <https://doi.org/10.1371/journal.pone.0111436>
- Kendall, W.L., Peterjohn, B.G., Sauer, J.R., 1996. First-Time Observer Effects in the North American Breeding Bird Survey. *The Auk* 113, 823–829. <https://doi.org/10.2307/4088860>
- Kéry, M., Royle, A., Meredith, M., 2023. AHMbook: Functions and Data for the Book “Applied Hierarchical Modeling in Ecology” Vols 1 and 2.
- Kéry, M., Royle, J.A., 2015. Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models, 1st edition. ed. Academic Press, Amsterdam ; Boston.
- Kéry, M., Royle, J.A., Plattner, M., Dorazio, R.M., 2009. Species richness and occupancy estimation in communities subject to temporary emigration. *Ecology* 90, 1279–1290. <https://doi.org/10.1890/07-1794.1>
- King, R., 2014. Statistical Ecology. *Annu. Rev. Stat. Its Appl.* 1, 401–426. <https://doi.org/10.1146/annurev-statistics-022513-115633>
- Lehikoinen, A., Lindström, Å., Santangeli, A., Sirkiä, P.M., Brotons, L., Devictor, V., Elts, J., Foppen, R.P.B., Heldbjerg, H., Herrando, S., Herremans, M., Hudson, M.-A.R., Jiguet, F., Johnston, A., Lorrilliere, R., Marjakangas, E.-L., Michel, N.L., Moshøj, C.M., Nellis, R., Paquet, J.-Y., Smith, A.C., Szép, T., van Turnhout, C., 2021. Wintering bird communities are tracking climate change faster than breeding communities. *J. Anim. Ecol.* 90, 1085–1095. <https://doi.org/10.1111/1365-2656.13433>
- Lehikoinen, E., Sparks, T.H., 2010. 9. Changes in migration, in: Møller, A.P., Fiedler, W., Berthold, P. (Eds.), *Effects of Climate Change on Birds*. Oxford University Press, Oxford ; New York.
- Lockwood, J.L., McKinney, M.L., 2001. Biotic Homogenization: A Sequential and Selective Process, in: Lockwood, J.L., McKinney, M.L. (Eds.), *Biotic Homogenization*. Springer US, Boston, MA. <https://doi.org/10.1007/978-1-4615-1261-5>
- Lorel, C., Le Viol, I., Plutzer, C., Jiguet, F., Mouchet, M., 2021. Linking the diversity and structure of French avian communities with landscape parameters, climate and NPP flows. *Reg. Environ. Change* 21, 61. <https://doi.org/10.1007/s10113-021-01786-y>
- Mackenzie, D.I., Nichols, James.D., Sutton, N., Kawanishi, K., Bailey, L.L., 2005. Improving Inferences in Population Studies of Rare Species That Are Detected Imperfectly. *Ecology* 86, 1101–1113.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., Langtimm, C.A., 2002. Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One. *Ecology* 83, 2248–2255. <https://doi.org/10.2307/3072056>

- Mathieu, J., Reynolds, J.W., Fragoso, C., Hadly, E., 2024. Multiple invasion routes have led to the pervasive introduction of earthworms in North America. *Nat. Ecol. Evol.* 8, 489–499. <https://doi.org/10.1038/s41559-023-02310-7>
- Maurer, B. a., 1998. The evolution of body size in birds. I. Evidence for non-random diversification. *Evol. Ecol.* 12, 925–934. <https://doi.org/10.1023/A:1006512121434>
- Maurer, B.A., Linder, E.T., Gammon, D., 2001. A Geographical Perspective on the Biotic Homogenization Process: Implications from the Macroecology of North American Birds, in: Lockwood, J.L., McKinney, M.L. (Eds.), *Biotic Homogenization*. Springer US, Boston, MA. <https://doi.org/10.1007/978-1-4615-1261-5>
- Maurer, B.A., McGill, B.J., 2010. Measurement of species diversity, in: Magurran, A.E., McGill, B.J. (Eds.), *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford University Press, Oxford, New York, pp. 55–64.
- Monnet, A.-C., Jiguet, F., Meynard, C.N., Mouillot, D., Mouquet, N., Thuiller, W., Devictor, V., 2014. Asynchrony of taxonomic, functional and phylogenetic diversity in birds. *Glob. Ecol. Biogeogr.* 23, 780–788. <https://doi.org/10.1111/geb.12179>
- Morelli, F., Brlík, V., Benedetti, Y., Bussière, R., Moudrá, L., Reif, J., Svitok, M., 2022. Detection Rate of Bird Species and What It Depends on: Tips for Field Surveys. *Front. Ecol. Evol.* 9. <https://doi.org/10.3389/fevo.2021.671492>
- Nabias, J., Barbaro, L., Fontaine, B., Dupuy, J., Couzi, L., Vallé, C., Lorrilliere, R., 2024. Reassessment of French breeding bird population sizes using citizen science and accounting for species detectability. *PeerJ* 12, e17889. <https://doi.org/10.7717/peerj.17889>
- Nolte, E., Bart, J., Pauli, B., Kaltenecker, G., Heath, J., 2016. Detectability of migrating raptors and its effect on bias and precision of trend estimates. *Avian Conserv. Ecol.* 11. <https://doi.org/10.5751/ACE-00894-110209>
- Noon, B.R., McKelvey, K.S., Lutz, D.W., LaHaye, W.S., Gutierrez, R.J., Moen, C.A., 1992. Estimates of demographic parameters and rates of population change. Chapter 8 Pages 175-186 *Calif. Spotted Owl Tech. Assess. Its Curr. Status Albany CA US Dep. Agric. For. Serv. Pac. Southwest Res. Stn. Gen Tech Rep PSW-GTR-133* 133, 186.
- Odom, K.J., Hall, M.L., Riebel, K., Omland, K.E., Langmore, N.E., 2014. Female song is widespread and ancestral in songbirds. *Nat. Commun.* 5, 3379. <https://doi.org/10.1038/ncomms4379>
- Olden, J., 2006. Biotic homogenization: A new research agenda for conservation biogeography. *J. Biogeogr.* 33, 2027–2039. <https://doi.org/10.1111/j.1365-2699.2006.01572.x>
- Pacifici, K., Reich, B.J., Miller, D.A.W., Gardner, B., Stauffer, G., Singh, S., McKerrow, A., Collazo, J.A., 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. *Ecology* 98, 840–850. <https://doi.org/10.1002/ecy.1710>
- Palacio, F.X., Maragliano, R.E., Montalti, D., 2020. The costs of ignoring species detectability on functional diversity estimation. *The Auk* 137, ukaa057. <https://doi.org/10.1093/auk/ukaa057>
- Pellissier, V., Touroult, J., Julliard, R., Sibley, J.P., Jiguet, F., 2013. Assessing the Natura 2000 network with a common breeding birds survey. *Anim. Conserv.* 16, 566–574. <https://doi.org/10.1111/acv.12030>

- Pilotto, F., Kühn, I., Adrian, R., Alber, R., Alignier, A., Andrews, C., Bäck, J., Barbaro, L., Beaumont, D., Beenaerts, N., Benham, S., Boukal, D.S., Bretagnolle, V., Camatti, E., Canullo, R., Cardoso, P.G., Ens, B.J., Everaert, G., Evtimova, V., Feuchtmayr, H., García-González, R., Gómez García, D., Grandin, U., Gutowski, J.M., Hadar, L., Halada, L., Halassy, M., Hummel, H., Huttunen, K.-L., Jaroszewicz, B., Jensen, T.C., Kalivoda, H., Schmidt, I.K., Kröncke, I., Leinonen, R., Martinho, F., Meesenburg, H., Meyer, J., Minerbi, S., Monteith, D., Nikolov, B.P., Oro, D., Ozoliņš, D., Padedda, B.M., Pallett, D., Pansera, M., Pardal, M.Â., Petriccione, B., Pipan, T., Pöyry, J., Schäfer, S.M., Schaub, M., Schneider, S.C., Skuja, A., Soetaert, K., Springe, G., Stanchev, R., Stockan, J.A., Stoll, S., Sundqvist, L., Thimonier, A., Van Hoey, G., Van Ryckegem, G., Visser, M.E., Vorhauser, S., Haase, P., 2020. Meta-analysis of multidecadal biodiversity trends in Europe. *Nat. Commun.* 11, 3486. <https://doi.org/10.1038/s41467-020-17171-y>
- Plummer, M., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proc. 3rd Int. Workshop Distrib. Stat. Comput.*
- Pollock, K.H., 1982. A Capture-Recapture Design Robust to Unequal Probability of Capture. *J. Wildl. Manag.* 46, 752–757. <https://doi.org/10.2307/3808568>
- Princé, K., Moussus, J.-P., Jiguet, F., 2012. Mixed effectiveness of French agri-environment schemes for nationwide farmland bird conservation. *Agric. Ecosyst. Environ.* 149, 74–79. <https://doi.org/10.1016/j.agee.2011.11.021>
- Princé, K., Zuckerberg, B., 2015. Climate change in our backyards: the reshuffling of North America's winter bird communities. *Glob. Change Biol.* 21, 572–585. <https://doi.org/10.1111/gcb.12740>
- Rao, C.R., 1982. Diversity and dissimilarity coefficients: A unified approach. *Theor. Popul. Biol.* 21, 24–43. [https://doi.org/10.1016/0040-5809\(82\)90004-1](https://doi.org/10.1016/0040-5809(82)90004-1)
- Richter, A., Nakamura, G., Agra Iserhard, C., da Silva Duarte, L., 2021. The hidden side of diversity: Effects of imperfect detection on multiple dimensions of biodiversity. *Ecol. Evol.* 11, 12508–12519. <https://doi.org/10.1002/ece3.7995>
- Ricotta, C., 2005. Additive partitioning of Rao's quadratic diversity: a hierarchical approach. *Ecol. Model.* 183, 365–371. <https://doi.org/10.1016/j.ecolmodel.2004.08.020>
- Riddle, J.D., Mordecai, R.S., Pollock, K.H., Simons, T.R., 2010. Effects of Prior Detections on Estimates of Detection Probability, Abundance, and Occupancy - Efectos de Detecciones Previas sobre Estimados de Probabilidad de Detección, Abundancia y Ocupación. *The Auk* 127, 94–99. <https://doi.org/10.1525/auk.2009.09062>
- Riebel, K., Odom, K.J., Langmore, N.E., Hall, M.L., 2019. New insights from female bird song: towards an integrated approach to studying male and female communication roles. *Biol. Lett.* 15, 20190059. <https://doi.org/10.1098/rsbl.2019.0059>
- Rigal, S., Devictor, V., Gaüzère, P., Kéfi, S., Forsman, J.T., Kajanus, M.H., Mönkkönen, M., Dakos, V., 2022. Biotic homogenisation in bird communities leads to large-scale changes in species associations. *Oikos* 2022, e08756. <https://doi.org/10.1111/oik.08756>
- Robinson, R.A., Julliard, R., Saracco, J.F., 2009. Constant effort: Studying avian population processes using standardised ringing. *Ringing Migr.* 24, 199–204. <https://doi.org/10.1080/03078698.2009.9674392>

- Rocchini, D., Marcantonio, M., Ricotta, C., 2017. Measuring Rao's  $Q$  diversity index from remote sensing: An open source solution. *Ecol. Indic.* 72, 234–238. <https://doi.org/10.1016/j.ecolind.2016.07.039>
- Rosenstock, S.S., Anderson, D.R., Giesen, K.M., Leukering, T., Carter, M.F., 2002. Landbird Counting Techniques: Current Practices and an Alternative. *The Auk* 119, 46–53. <https://doi.org/10.1093/auk/119.1.46>
- Ross, C., Reeve, N., 2011. Survey and census methods: population distribution and density, in: Setchell, J.M., Curtis, D.J. (Eds.), *Field and Laboratory Methods in Primatology: A Practical Guide*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511921643>
- Roth, T., Allan, E., Pearman, P.B., Amrhein, V., 2018. Functional ecology and imperfect detection of species. *Methods Ecol. Evol.* 9, 917–928. <https://doi.org/10.1111/2041-210X.12950>
- Schaub, M., Abadi, F., 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. *J. Ornithol.* 152, 227–237. <https://doi.org/10.1007/s10336-010-0632-7>
- Schmidt, B.R., 2005. Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 15, 681–692. <https://doi.org/10.1002/aqc.740>
- Sheard, C., Neate-Clegg, M.H.C., Alioravainen, N., Jones, S.E.I., Vincent, C., MacGregor, H.E.A., Bregman, T.P., Claramunt, S., Tobias, J.A., 2020. Ecological drivers of global gradients in avian dispersal inferred from wing morphology. *Nat. Commun.* 11, 2463. <https://doi.org/10.1038/s41467-020-16313-6>
- Si, X., Cadotte, M.W., Zhao, Y., Zhou, H., Zeng, D., Li, J., Jin, T., Ren, P., Wang, Y., Ding, P., Tingley, M.W., 2018. The importance of accounting for imperfect detection when estimating functional and phylogenetic community structure. *Ecology* 99, 2103–2112. <https://doi.org/10.1002/ecy.2438>
- Sólymos, P., Matsuoka, S.M., Stralberg, D., Barker, N.K.S., Bayne, E.M., 2018. Phylogeny and species traits predict bird detectability. *Ecography* 41, 1595–1603. <https://doi.org/10.1111/ecog.03415>
- Strebel, N., Kéry, M., Schaub, M., Schmid, H., 2014. Studying phenology by flexible modelling of seasonal detectability peaks. *Methods Ecol. Evol.* 5, 483–490. <https://doi.org/10.1111/2041-210X.12175>
- Székely, T., Lislevand, T., Figuerola, J., 2007. Sexual size dimorphism in birds, in: *Sex, Size and Gender Roles*. pp. 27–37. <https://doi.org/10.1093/acprof:oso/9780199208784.003.0004>
- Thompson, W.L., 2002. Towards Reliable Bird Surveys: Accounting for Individuals Present but not Detected. *The Auk* 119, 18–25. [https://doi.org/10.1642/0004-8038\(2002\)119\[0018:TRBSAF\]2.0.CO;2](https://doi.org/10.1642/0004-8038(2002)119[0018:TRBSAF]2.0.CO;2)
- Tilman, D., 2001. Functional Diversity, in: Levin, S.A. (Ed.), *Encyclopedia of Biodiversity*. Academic Press, pp. 109–120.
- Tilman, D., Reich, P.B., Knops, J., Wedin, D., Mielke, T., Lehman, C., 2001. Diversity and Productivity in a Long-Term Grassland Experiment. *Science* 294, 843–845. <https://doi.org/10.1126/science.1060391>

- Tingley, M.W., Beissinger, S.R., 2013. Cryptic loss of montane avian richness and high community turnover over 100 years. *Ecology* 94, 598–609. <https://doi.org/10.1890/12-0928.1>
- Tingley, M.W., Nadeau, C.P., Sandor, M.E., 2020. Multi-species occupancy models as robust estimators of community richness. *Methods Ecol. Evol.* 11, 633–642. <https://doi.org/10.1111/2041-210X.13378>
- Tobias, J.A., Sheard, C., Pigot, A.L., Devenish, A.J.M., Yang, J., Sayol, F., Neate-Clegg, M.H.C., Alioravainen, N., Weeks, T.L., Barber, R.A., Walkden, P.A., MacGregor, H.E.A., Jones, S.E.I., Vincent, C., Phillips, A.G., Marples, N.M., Montaña-Centellas, F.A., Leandro-Silva, V., Claramunt, S., Darski, B., Freeman, B.G., Bregman, T.P., Cooney, C.R., Hughes, E.C., Capp, E.J.R., Varley, Z.K., Friedman, N.R., Korntheuer, H., Corrales-Vargas, A., Trisos, C.H., Weeks, B.C., Hanz, D.M., Töpfer, T., Bravo, G.A., Remeš, V., Nowak, L., Carneiro, L.S., Moncada R., A.J., Matysioková, B., Baldassarre, D.T., Martínez-Salinas, A., Wolfe, J.D., Chapman, P.M., Daly, B.G., Sorensen, M.C., Neu, A., Ford, M.A., Mayhew, R.J., Fabio Silveira, L., Kelly, D.J., Annorbah, N.N.D., Pollock, H.S., Grabowska-Zhang, A.M., McEntee, J.P., Carlos T. Gonzalez, J., Meneses, C.G., Muñoz, M.C., Powell, L.L., Jamie, G.A., Matthews, T.J., Johnson, O., Brito, G.R.R., Zyskowski, K., Crates, R., Harvey, M.G., Jurado Zevallos, M., Hosner, P.A., Bradfer-Lawrence, T., Maley, J.M., Stiles, F.G., Lima, H.S., Provost, K.L., Chibesa, M., Mashao, M., Howard, J.T., Mlamba, E., Chua, M.A.H., Li, B., Gómez, M.I., García, N.C., Päckert, M., Fuchs, J., Ali, J.R., Derryberry, E.P., Carlson, M.L., Urriza, R.C., Brzeski, K.E., Prawiradilaga, D.M., Rayner, M.J., Miller, E.T., Bowie, R.C.K., Lafontaine, R.-M., Scofield, R.P., Lou, Y., Somarathna, L., Lepage, D., Illif, M., Neuschulz, E.L., Templin, M., Dehling, D.M., Cooper, J.C., Pauwels, O.S.G., Analuddin, K., Fjeldså, J., Seddon, N., Sweet, P.R., DeClerck, F.A.J., Naka, L.N., Brawn, J.D., Aleixo, A., Böhning-Gaese, K., Rahbek, C., Fritz, S.A., Thomas, G.H., Schleuning, M., 2022. AVONET: morphological, ecological and geographical data for all birds. *Ecol. Lett.* 25, 581–597. <https://doi.org/10.1111/ele.13898>
- Vallé, C., Le Viol, I., Kerbirriou, C., Bas, Y., Jiguet, F., Princé, K., 2023. Farmland biodiversity benefits from small woody features. *Biol. Conserv.* 286, 110262. <https://doi.org/10.1016/j.biocon.2023.110262>
- Villéger, S., Mason, N.W.H., Mouillot, D., 2008. New Multidimensional Functional Diversity Indices for a Multifaceted Framework in Functional Ecology. *Ecology* 89, 2290–2301. <https://doi.org/10.1890/07-1206.1>
- Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E., 2007. Let the concept of trait be functional! *Oikos* 116, 882–892. <https://doi.org/10.1111/j.0030-1299.2007.15559.x>
- Yoccoz, N.G., Nichols, J.D., Boulinier, T., 2001. Monitoring of biological diversity in space and time. *Trends Ecol. Evol.* 16, 446–453. [https://doi.org/10.1016/S0169-5347\(01\)02205-4](https://doi.org/10.1016/S0169-5347(01)02205-4)
- Zipkin, E.F., Zylstra, E.R., Wright, A.D., Saunders, S.P., Finley, A.O., Dietze, M.C., Itter, M.S., Tingley, M.W., 2021. Addressing data integration challenges to link ecological

*Chapitre 3 : Effet de la non prise en compte du processus d'observation dans l'estimation d'indice de biodiversité*

processes across scales. Front. Ecol. Environ. 19, 30–38.  
<https://doi.org/10.1002/fee.2290>



## VI Discussion générale

---

### VI.A Rappel des résultats

La quantité d'individus et leur distribution spatiale constituent des informations essentielles pour la surveillance des populations. L'observation directe ne révèle qu'une portion limitée de l'information sur l'abondance des populations déterminée par un ensemble de dynamique complexe, le processus écologique. L'estimation d'indicateurs d'abondance non biaisés nécessite la prise en compte du processus d'observation déterminant les variations du nombre d'individus détectés lors de la collecte.

Au cours de cette thèse, nos recherches se sont d'abord concentrées sur l'estimation des effectifs d'oiseaux nicheurs. Dans le cadre de l'actualisation des effectifs d'oiseaux nicheurs communs (chapitre 1), nous avons utilisé des données de sciences participatives issues d'un protocole structuré (EPOC-ODF) ainsi que des modèles hiérarchiques afin d'estimer le nombre d'individus tout en prenant en compte les variations liées au processus d'observation. La comparaison des nouvelles estimations avec celles du précédent atlas, tout en tenant compte des tendances populationnelles, a révélé une potentielle sous-estimation généralisée des espèces communes nicheuses de France métropolitaine lors des précédents atlas (voir le tableau des estimations : *Annexe: Tableau des tailles de populations d'oiseaux communs nicheurs de France métropolitaine*).

L'estimation non biaisée d'abondance requiert une information structurée demandant un effort de collecte accru. Dans le contexte des sciences participatives, basées sur le volontariat des observateurs, cela induit un compromis entre (i) la collecte d'une information très qualitative, mais en faible quantité, et (ii) la collecte d'une information plus abondante, mais de qualité moindre. Pour répondre à ce dilemme, nous avons mis en place un modèle intégré (chapitre 2) qui utilise simultanément les données collectées par les protocoles EPOC-

ODF et EPOC dans l'inférence d'abondance. Nos travaux ont mis en évidence que l'intégration des deux jeux de données permet de réduire l'incertitude liée à l'inférence du processus d'observation et du processus écologique. Une étude de cas réalisée sur des données réelles démontre que cette réduction d'incertitude dépend du ratio entre les types de données utilisées.

L'abondance correspond à une information élémentaire utilisée pour décrire des aspects plus complexes des communautés écologiques. Dans ce dernier chapitre (chapitre 3), nous nous sommes concentrés sur l'estimation des indices de communautés. Nous avons démontré que l'omission du processus d'observation, lors de l'estimation d'abondance des espèces, se répercute sur les indices de communautés entraînant des biais dans l'évaluation des diversités taxonomique, phylogénétique et fonctionnelle.

Dans la discussion qui suit, je compte aborder les perspectives de nos travaux, notamment à la suite du chapitre 2, afin d'estimer plus en détail les dynamiques des populations des oiseaux nicheurs de France métropolitaine via la mise en place de modèles de population intégrés. Les travaux réalisés durant cette thèse abordent principalement l'aspect numérique des indices de biodiversité. Dans un second temps, je reviens et discute sur la notion d'indice en intégrant la composante sociale et l'intérêt de sa bonne diffusion. Pour illustrer ce point, je prends comme exemple l'indice planète vivante (LPI) pour illustrer les potentielles tensions entre l'objectif initial de l'indice, son interprétation et sa portée de diffusion. Enfin, je contextualise la production d'indices notamment leur rôle dans la prise de décision et la place de l'incertitude scientifique.

## VI.B Estimation d'abondance : vers une vision plus détaillée de l'avifaune française

Les résultats de nos recherches corroborent un ensemble de recherches sur l'utilité des sciences participatives pour l'acquisition de données écologiques et l'évaluation des états de population de façon directe (Devictor, Whittaker & Beltrame, 2010; McKinley et al., 2017; Chandler et al., 2017; Fink et al., 2020; Fontaine et al., 2022) ou indirecte (Matutini et al., 2021) malgré des biais (Isaac et al., 2014) adressables en amont de la collecte ou durant l'analyse des données (Kosmala et al., 2016; Johnston et al., 2019; Callaghan et al., 2019; Johnston, Matechou & Dennis, 2022). Cependant, dans l'ensemble des travaux que nous avons réalisés, nous avons soit restreint la période temporelle à un intervalle de trois ans (chapitre 1 et 2), soit considéré une tendance linéaire et homogène sur l'ensemble de la population (chapitre 3). Bien que nous assumions ces approximations et simplifications, nous reconnaissons que les dynamiques des populations sont plus complexes (Rigal, Devictor & Dakos, 2020) et que les changements globaux peuvent avoir un effet hétérogène sur l'ensemble de celles-ci, *i.e.* mâles, femelles, adultes et juvéniles (Van De Pol et al., 2010; Moreno & Møller, 2011; Kissel et al., 2019). Une population peut être représentée par le cycle de vie des individus qui la composent où les paramètres démographiques (fécondité, survie, longévité, immigration et émigration) influent sur la quantité d'individus à chaque stade (Figure 28; Legendre, 2020).

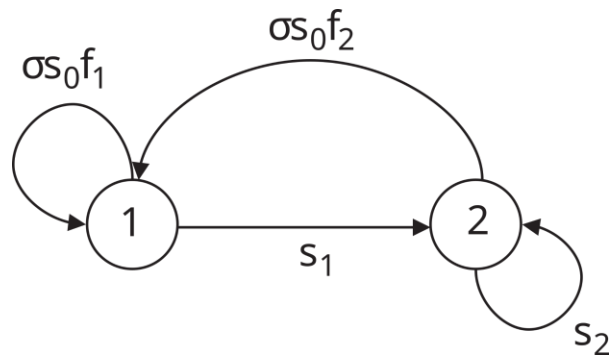


Figure 28 : Exemple du cycle de vie d'un passereau (selon un décompte avant reproduction) selon deux stades de vie ; 1 : individu de 1 an (sous-adulte) et 2 : individu de plus de deux ans (adulte). Chaque stade de vie possède un taux de survie et de fécondité propre, respectivement ( $S_1$ ,  $S_2$ ,  $F_1$  et  $F_2$ ). Comme le décompte de la population se fait en amont de la reproduction, les nouveau-nés sont comptés l'année suivante et possèdent un taux de survie  $S_0$ . Ce cycle de vie considère uniquement les femelles nécessitant l'application d'un paramètre qui décrit la répartition de femelles dans la population, le sex-ratio. Figure tirée de Legendre 2020.

L'estimation des différents paramètres démographiques nécessite un ensemble de données spécifiques, notamment des données de baguage (Robinson, Julliard & Saracco, 2009) pour l'estimation des taux de survie, et les données de comptage pour l'estimation des quantités d'individus (Emlen, 1971). L'analyse conjointe de ces différentes sources de données est réalisable à l'aide de modèles de population intégrés – *Integrated Population Models* (IPM ; Figure 29; Besbeas et al., 2002; Schaub & Abadi, 2010; Quéroué et al., 2021) qui permet d'estimer avec une plus grande précision les paramètres démographiques (Abadi et al., 2010) tout en tenant compte de leurs incertitudes respectives (Schaub, 2020). En considérant l'ensemble des programmes de suivis de l'avifaune mis en place sur le territoire français pour le décompte d'oiseaux (EPOC, EPOC-ODF, STOC-EPS, SHOC<sup>34</sup>, STOM<sup>35</sup>) et le baguage (STOC-Capture), il serait possible de mettre en place un suivi intégré des populations de l'avifaune française (Julliard & Jiguet, 2002).

<sup>34</sup> SHOC : Suivi Hivernal des Oiseaux Communs (<https://www.oiseauxdefrance.org/get-involved/shoc>)

<sup>35</sup> STOM : Suivi Temporel des Oiseaux de Montagnes (<https://www.lpo.fr/la-lpo-en-actions/connaissance-des-especes-sauvages/suivis-ornithologiques/oiseaux-communs/stom>)

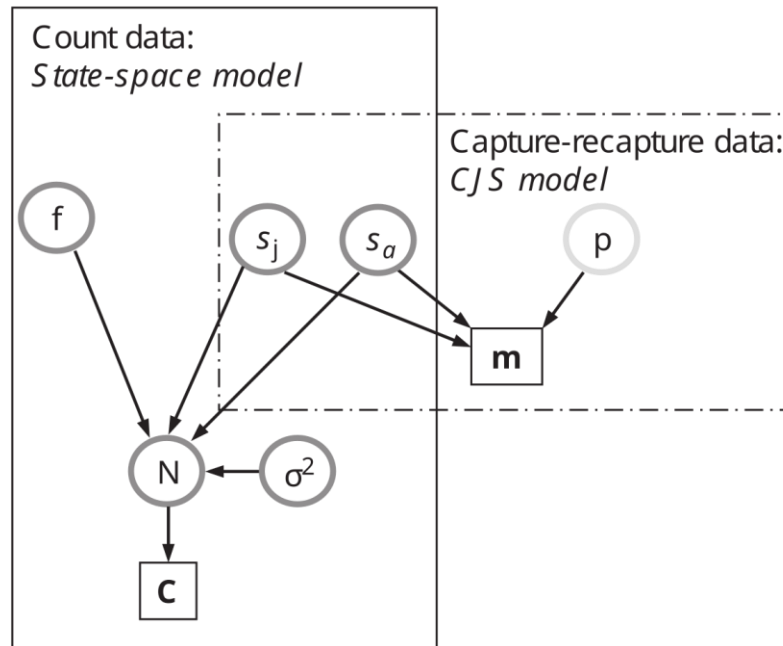


Figure 29 : Graphique d'un exemple d'IPM représentant les liens entre les différents paramètres estimés : paramètres désirés (cercles gris), paramètres de nuisance (cercles blancs) et données collectées (carrés noirs). Les carrés représentent les données collectées. De haut en bas et gauche à droite :  $f$  : fécondité (ou productivité) ;  $s_j$  : survie des juvéniles ;  $s_a$  : survie des adultes ;  $p$  : probabilité de recapture ;  $m$  : données de capture-recapture ;  $N$  : taille de population estimée (totale ou par stades de vie) ;  $\sigma^2$  : erreur d'observation ;  $C$  : donnée de comptage. Les sous-modèles qui composent l'IPM sont représentés par des rectangles. Tiré de Schaub, 2020.

## VI.C Indices et narratifs : une association nécessaire

‘Indicateur = indice + narratif’ – Séminaire du groupe indicateur CESCO (2023)

‘How is the observation of the world communicated within the world?’<sup>36</sup> – Luhmann, 1998, The Ecology of Ignorance. In Observation on modernity.

### VI.C.i Indicateurs : un ensemble de critères et de tensions

Les indicateurs ne sont pas uniquement constitués de connaissances scientifiques, mais peuvent aussi dépendre de composantes sociales (Ladle et al., 2016; Reyes-García et al., 2022). Par exemple, une espèce convoitée devient d’autant plus désirée par les braconniers en fonction de sa rareté – *l’effet Allee anthropogénique* (Courchamp et al., 2006) – tandis que des espèces d’une grande importance culturelle sont plus susceptibles de bénéficier d’actions de conservation importantes ou de stratégies de gestion durable (Figure 30; Ladle et al., 2023).

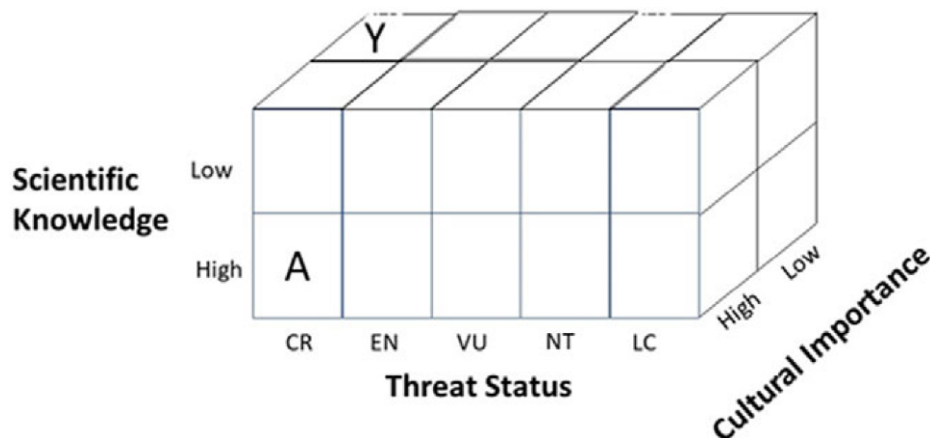


Figure 30 : Proposition d’amélioration des critères de menace des listes rouges de l’IUCN intégrant simultanément les connaissances scientifiques et l’importance culturelle des espèces. Les espèces présentes dans le bloc Y sont plus enclines à s’éteindre que les espèces du bloc A. En effet, en plus de facteurs biologiques d’extinction (e.g., faible quantité d’individus, aire de distribution restreinte, effectifs déclinants ...), les espèces du bloc Y ont une (i) une faible

<sup>36</sup> Citation faite dans le contexte de la sociologie. ‘Sociology must again modify its theory of a self-observing world and pose the question...’.

importance culturelle réduisant la volonté d'agir, et (ii) peu de connaissances scientifiques permettant la mise en place d'actions efficaces. Tiré de Ladle et al., 2023

Ces indicateurs ont pour principal objectif de synthétiser une information écologique afin de faciliter sa transmission et d'assurer son suivi et sa gestion (Levrel, 2007). Pour assurer une utilisation optimale par l'ensemble des acteurs, ils doivent respecter six critères (Desrosières, 2003) à savoir:

- (i) la pertinence : ils doivent être en adéquation avec les besoins des acteurs;
- (ii) la précision : la valeur décrite par l'indice doit être proche de la valeur réelle;
- (iii) l'actualité et la ponctualité vis-à-vis des échéances décisionnelles ou de reportages;
- (iv) l'accessibilité et la clarté lors de l'utilisation ou leur interprétation;
- (v) la comparabilité des données;
- (vi) la cohérence entre l'indice et son interprétation.

Cependant, il est impossible de répondre complètement à l'ensemble des critères (Figure 31), il existe des tensions notamment en fonction de différents gradients. Un indicateur contextuel comme celui des espèces indicatrices a pour objectif de synthétiser la réponse d'un habitat ou d'une communauté face à des changements, mais tend à omettre la réponse des espèces non ciblées, par exemple le suivi des abeilles met en évidence les effets néfastes des pratiques agricoles (Quigley, Amdam & Harwood, 2019). À l'inverse, un indice universel et extensible comme l'indice de planète vivante<sup>37</sup> peut synthétiser les variations temporelles d'un ensemble de population selon plusieurs échelles spatiales (Ledger et al., 2023), mais cela peut entraîner une perte du signal des variations spécifiques.

---

<sup>37</sup> On aborde plus en détails cet indice dans la section suivante

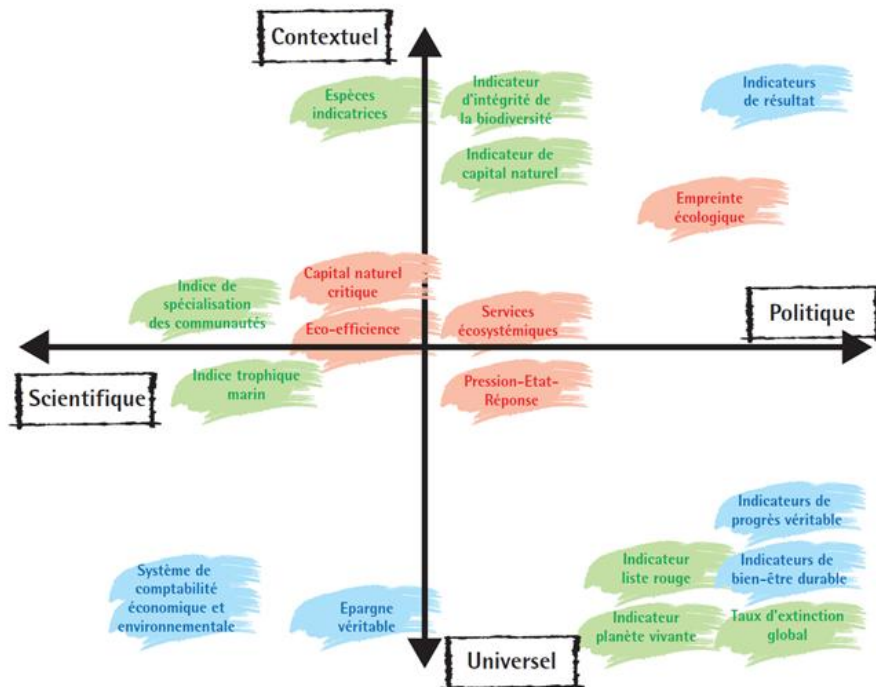


Figure 31 : Distribution d'un ensemble d'indicateurs selon leurs dimensions scientifique et politique d'une part, universelle et contextuelle de l'autre. Ces indicateurs peuvent être classés en fonction de différentes approches : en vert, ceux qui reposent sur une perspective écolo-centrée ; en bleu, ceux qui reflètent une approche économico-centrée ; et en rouge, ceux qui adoptent une vision plus transdisciplinaire. Tiré de Levrel et al., 2007

Un autre écart que peuvent avoir les indicateurs correspond à leur portée, l'idée étant de construire et diffuser l'information de la manière la plus efficace. Bien que la construction d'indices soit au cœur de plusieurs disciplines scientifiques, *i.e.* biologie de la conservation et biogéographie de la conservation (Whittaker & Ladle, 2011), les scientifiques ne sont pas forcément formés pour aborder leur diffusion (Brownell, Price & Steinman, 2013; Cooke et al., 2017). En se basant sur ces critères, on constate que, dans le cadre du suivi de l'avifaune française, les points (i) et (iii) sont respectés par l'exercice du rapportage de la Directive 'Oiseaux' tous les six ans. Les résultats des chapitres 1 et 2 montrent que les sciences citoyennes permettent d'obtenir des estimations plus fiables, précises (lors de l'intégration des données). Même si les résultats, sous la forme de fourchette de population ou de carte de distribution spatiale, sont accessibles (iv), le cheminement analytique reste difficile à saisir.

C'est principalement dû à l'écart entre la donnée collectée par les participants, à savoir le pointage de localisation d'individus (EPOC et EPOC-ODF) ou la détermination d'intervalles de distances (STOC-EPS) et l'information délivrée. Il est nécessaire de justifier auprès des observateurs les raisons de l'effort supplémentaire demandé lors de la collecte, d'un point de vue éthique, mais aussi dans un objectif de dissémination de l'approche scientifique<sup>38</sup> (Bonney et al., 2009). Au sujet de l'échelle de l'indice en amont (Figure 31), les estimations d'abondances ont pour objectif d'informer sur la taille de populations nationales, mais il se peut que certaines estimations à des échelles régionales/communales diffèrent, dans la mesure où la résolution de l'échantillonnage (2 x 2km) peut ne pas suffire pour représenter fidèlement leurs habitats (Wiens, 1989; *Spécificité du territoire français métropolitain*).

#### VI.C.ii L'agrégation d'indices multi-espèces : cas du LPI

On peut différencier deux types d'indicateurs (i) les indicateurs uniques basés sur une seule unité de mesure, comme l'individu, le gène ou l'interaction, et (ii) les indicateurs composites qui combinent plusieurs unités de mesure. Alors que les premiers doivent se confronter aux dilemmes mentionnés plus tôt, les seconds doivent également gérer les modalités de regroupement des différentes mesures et les variations entre celles-ci (Levrel, 2007). Au regard de ce point et des discussions précédentes, il est intéressant d'étudier plus en détail le cas du Living Planet Index (LPI; Collen et al., 2009). Il s'agit d'un indice grandement médiatisé, utilisé comme l'un des indicateurs de la Convention sur la Biologie et la Diversité (CBD ; Almond et al., 2022). Il utilise une information spécifique, *i.e.* les tendances d'abondance de populations, à une échelle contextuelle variable via des

---

<sup>38</sup> Il ne s'agit pas forcément d'un acte de bienveillance de la part des scientifiques. C'est aussi un moyen afin de transmettre ses valeurs, encadrées par la rigueur scientifique, d'affirmer sa position (financement d'étude et poursuite de carrière) et son utilité et sa pertinence au sein de la société (Beck et al., 2019).

regroupements spatiaux, taxonomiques ou fonctionnels<sup>39</sup> (Figure 32; Collen et al., 2009; Almond et al., 2022).

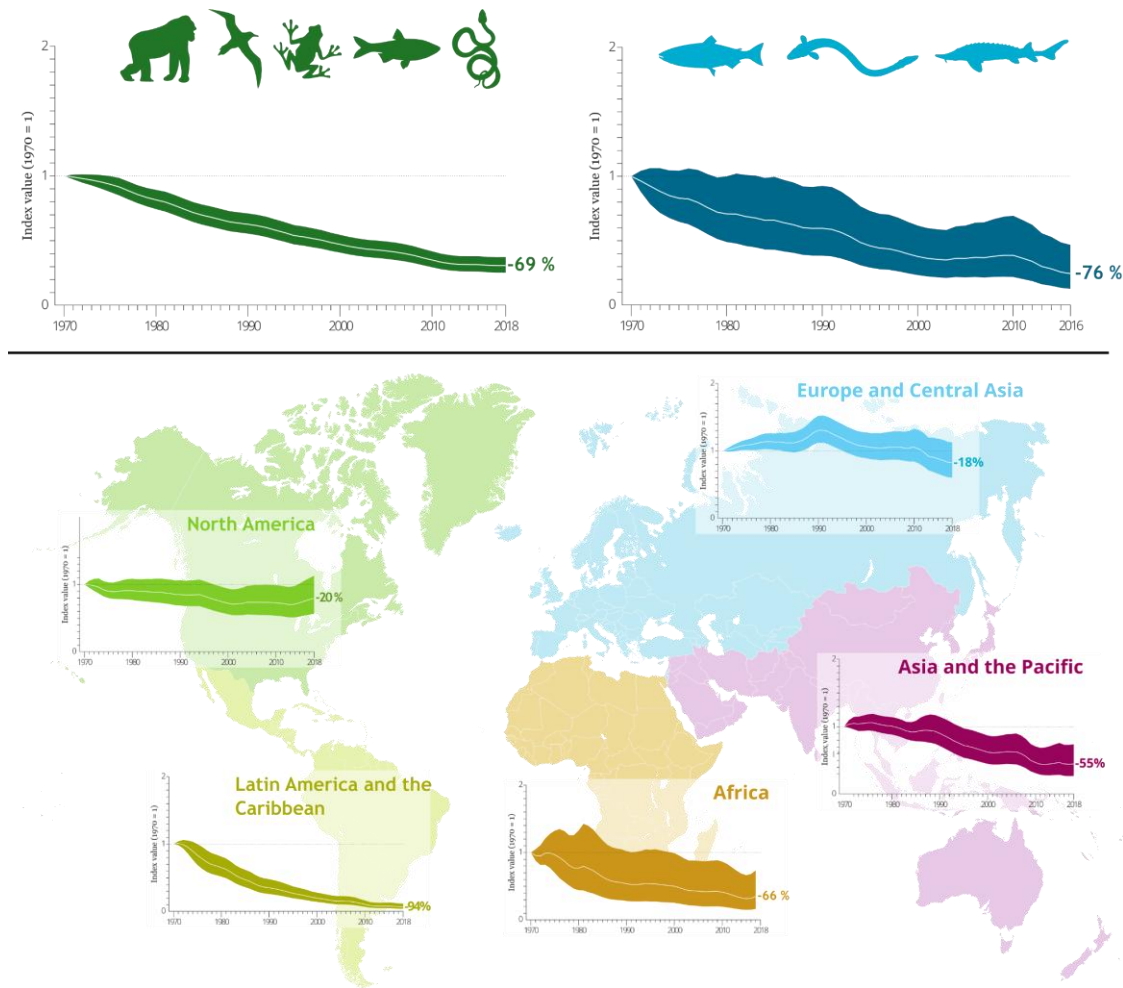


Figure 32: Représentation du LPI selon différentes échelles fonctionnelles et spatiales. De haut en bas et gauche à droite : LPI (1970-2018) global basé sur les variations d’abondances relatives de 31821 populations de 5230 espèces ; LPI (1970-2016) des poissons d’eaux douces migrateurs basé sur 1406 populations de 247 espèces ; LPI selon les régions de la plateforme intergouvernementale scientifique et politique sur la biodiversité et les services écosystémiques (IPBES ; [IPBES Technical Support Unit on Knowledge and, 2021](#)). Sur l’ensemble des figures, la tendance moyenne est décrite par une courbe blanche associée à un intervalle de confiance. Adapté de Almond et al., 2022.

Le LPI se calcule à l’aide des tendances temporelles de multiples populations à partir d’une année de référence (1970), son calcul se décompose de la manière suivante ([Collen et al., 2009](#); [Puurtinen, Elo & Kotiaho, 2022](#)) :

<sup>39</sup> En plus de l’exemple des poissons migrateurs, on peut noter les tendances temporelles des oiseaux communs, où les espèces sont regroupées selon leurs spécialisations à l’habitat (forestier, agricole, urbain ou généraliste). Voir la figure ‘Evolution des indicateurs par groupe de spécialisation’ de [Fontaine et al., 2020](#).

$$d_{it} = \log_{10}(N_{it}/N_{i(t-1)})$$

$$\bar{d}_t = 1/n_t \sum_{i=1}^{n_t} d_{it}$$

$$I_t = I_{t-1} * 10^{\bar{d}_t}$$

***i* : Espèces**

***t* : Années**

**$N_{i,t}$  : Tailles de populations mesurées de l'espèce *i* à l'année *t***

**$n_t$  : Nombre de populations à l'année *t***

**$d_{i,t}$  : Variations d'abondances de la population  $N_{i,t}$**

**$\bar{d}_t$  : Moyenne arithmétique des variations d'abondances à l'année *t***

**$I_t$  : Valeur du LPI à l'année *t***

Les valeurs du LPI en fonction de la zone géographique ou du regroupement, sont calculées en sous-échantillonnant les données populationnelles (Figure 32) ou en pondérant la moyenne arithmétique ( $\bar{d}_t$ ) en fonction de la proportion d'espèces de groupe (McRae, Deinet & Freeman, 2017). Bien que le LPI touche un large spectre d'acteurs allant des décideurs (Ledger et al., 2023) au grand public (Mouterde, 2020; Almond et al., 2022), son calcul et son interprétation sont controversés sur différents aspects, à savoir ; (i) le biais taxonomique lors de l'agrégation (McRae, Deinet & Freeman, 2017) ; (ii) l'impossibilité d'inclure des populations éteintes, ainsi qu'une sensibilité au déclin précoce, proche de 1970 (Toszygova, Smyčka & Storch, 2024) ; (iii) la divergence dans la qualité des données acquises (Figure 33; Marconi et al., 2021; Dove et al., 2023) ; (iv) l'interprétation non évidente de l'indicateur (Puurinen, Elo & Kotiaho, 2022; Ledger et al., 2023).

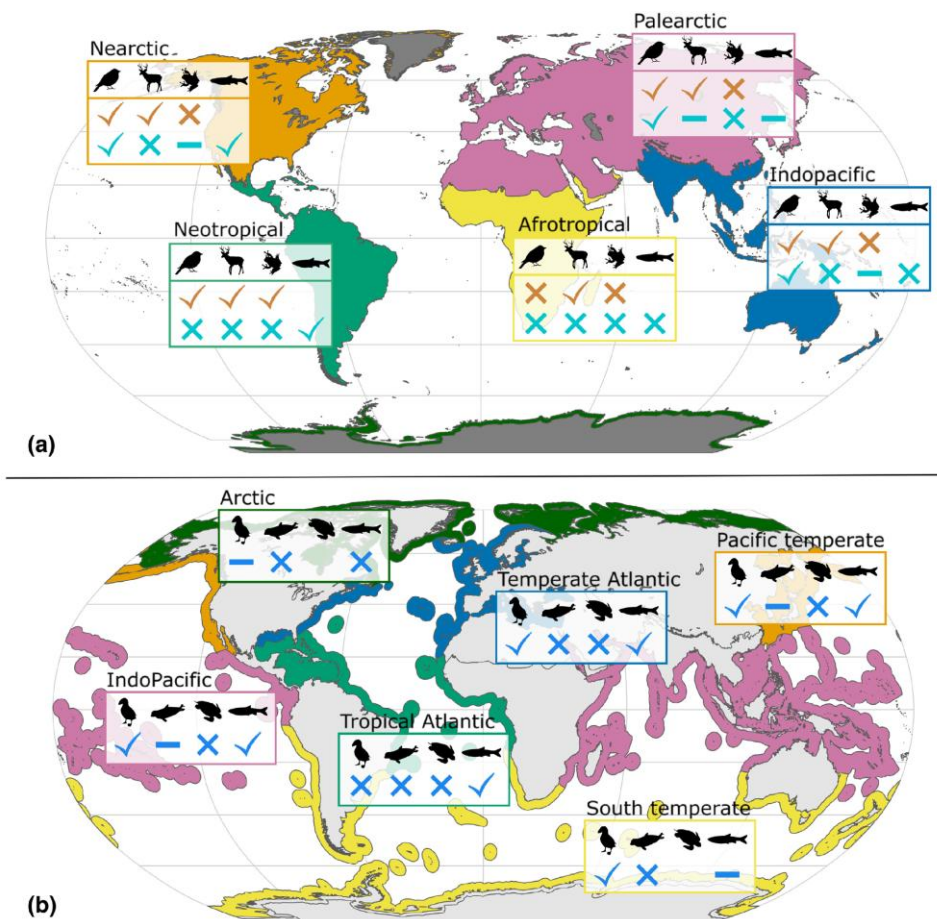


Figure 33 : Fiabilité des données utilisées pour calculer le LPI en fonction des régions de l'IPBES et des groupes taxonomiques. En (a) les animaux terrestres (en haut) et d'eaux douces (en bas) ; et en (b) les animaux marins. La fiabilité est déterminée par la quantité de données disponibles lors du calcul du LPI. Les groupes taxonomiques et zones IPBES avec une coche possèdent le nombre nécessaire (voire plus) de données pour le calcul du LPI ; ceux représentés par un tiret sont déficitaires (50% à 99%) de la quantité de données requises ; et ceux décrits par une croix sont sévèrement déficitaires (<50%). Tiré de Dove et al., 2023

Les trois premiers points abordés sont adressables par la mise en place de moyens de standardisation de la collecte de données<sup>40</sup> (Ledger et al., 2023), alors que le dernier point reflète un enjeu de taille pour le LPI lié à la volonté d'agréger des données relatives. Bien que les tendances populationnelles puissent être correctement estimées par l'acquisition de données relatives (Thompson, 2002; résultats de simulation de l'article 3), le manque d'intégration du processus de détection ne permet pas d'estimer l'information écologique

<sup>40</sup> Notamment le respect du principes FAIR – Findable, Accessible, Interoperable and Reusable ;(Wilkinson et al., 2016)

attendue, voire espérée, que l'indice vise à diffuser, *i.e. la variation multi-échelles, taxonomique et/ou fonctionnelles de l'abondance*. Le recours à des modèles hiérarchiques semble prometteur pour prendre en compte les soucis, liés à l'analyse et l'agrégation des données, décrits plus tôt ([Daskalova, Myers-Smith & Godlee, 2020](#)). Les recherches réalisées durant cette thèse semblent concourir aux besoins de plusieurs acteurs sur la mise en place d'outils permettant d'évaluer l'abondance de plusieurs espèces à partir de plusieurs jeux de données (Chapitres 2 et 3).

RESTRICTED TO SCIENTISTS

International Commission for



the Northwest Atlantic Fisheries

Serial No. 000<sup>1</sup>/<sub>2</sub>  
(WOW)

ICNAF Dumm.Doc. 75/2

ANNUAL MEETING - JUNE 1975

Estimation of Unknown Natural Mortality

by

J.G. Pope

$$M = ?$$

$$M = ?$$

$$M = .2$$

$$M = .2$$

$$M = 0.2$$

EUREKA!!!

Illustration : International Commission for the Northwest Atlantic Fisheries (ICNAF) 'Dummy Documents' by John Pope drawn at the April 1975 meeting of the ICNAF Assessment Subcommittee entitled "Estimation of Unknown Natural Mortality".

## VI.D Transmission de l'incertitude scientifique face à la prise d'action

“‘There is no evidence’ despite ‘over 20 years of intense study that neonicotinoids have any link to colony honeybee losses’ [...] The commission ‘responded to a ‘weight of noise’ instead of the ‘weight of evidence’ which clearly shows that in reality neonicotinoids pose a minimal threat to bee health compared with a lack of food, diseases, and cold weather’ – Bodoni, 2017. ‘Bayer, Syngenta Clash With EU Over Bees Amid M&A Charm Offensive’

‘For instance, a precautionary regime may require rigorous scientific evidence of low risk before an activity is allowed to proceed, and a non-precautionary regime may require rigorous scientific evidence of risk before prohibiting an activity. Which of these is chosen does not reflect differing valuations of science, but differing value judgements about what objectives should be favored when science is uncertain.’ – Cooney, 2005, 2. The meaning of the precautionary principle, In ‘Biodiversity and the Precautionary Principle: Risk and Uncertainty in Conservation and Sustainable Use’.

L'incertitude peut avoir différentes origines (Mace & Sissenwine, 2002), qui peuvent être (i) quantitatives, *i.e.* liées aux variations aléatoires du processus écologique et/ou du processus d'observation (Cf. *Le vivant caché par l'observation*) ou correspondre à un écart entre le modèle et la réalité (Figure 25A) ; mais aussi (ii) sociales lorsqu'il y a une différence entre les résultats scientifiques et les mesures employées par des acteurs politique, social ou financier (Maxim & Van Der Sluijs, 2007).

Lors d'une étude scientifique, il est difficile d'attribuer avec certitude l'importance de l'effet étudié<sup>41</sup> et de différencier les variations liées au processus, ou les variations résiduelles

---

<sup>41</sup> On peut prendre par exemple la figure 5 d'Efron, 2012, qui montre deux intervalles de confiance possibles incluant le même paramètre estimé  $\theta$

## Discussion générale Transmission de l'incertitude scientifique face à la prise d'action

liées à la collecte de données ou l'analyse (Norton, 1994). Bien que les modèles hiérarchiques permettent de prendre en compte le processus d'observation, ils constituent des représentations (plus complexes ; voir : *Le vivant caché par l'observation*) encore à la merci des décisions analytiques des écologues. On peut prendre en exemple l'étude préliminaire récente de Gould et al., 2023, montrant des résultats hétérogènes de l'effet de la taille des couvées sur la croissance des oisillons (Figure 34) obtenues par plusieurs équipes de recherches et reflète le besoin de transparence et le recours à la science ouverte afin de pouvoir nuancer et retracer les décisions prises (Nosek et al., 2015; Parker et al., 2016).

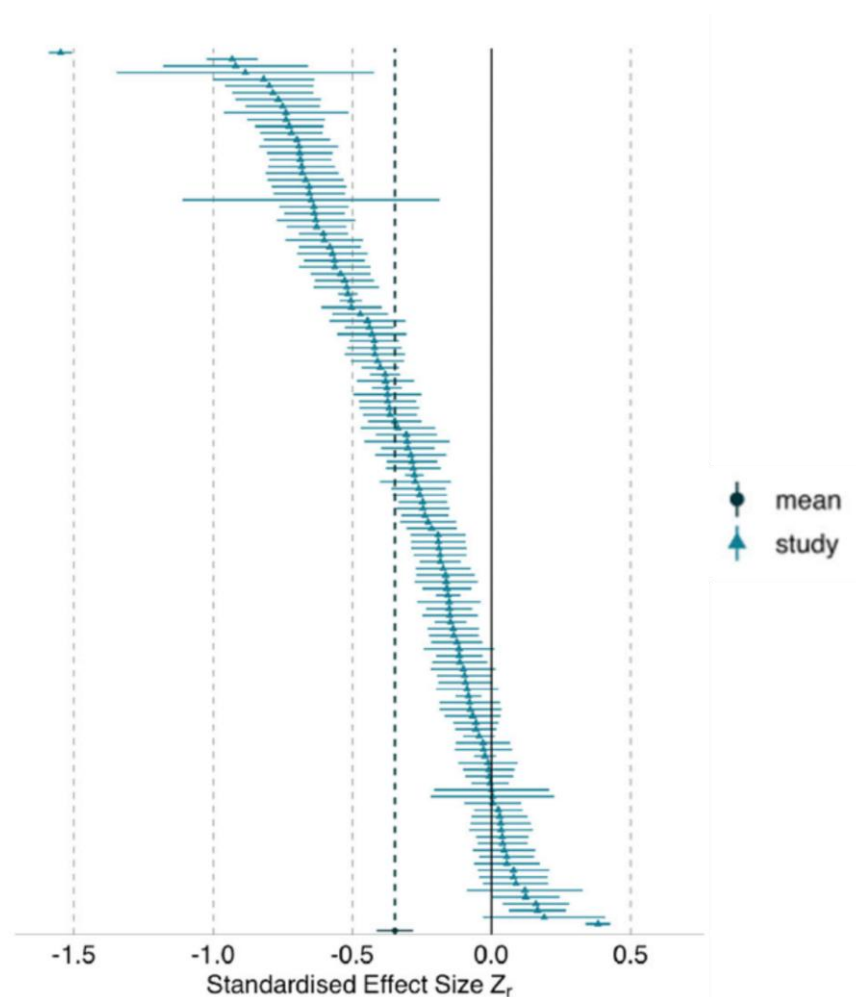


Figure 34 : Distribution de l'effet de la croissance des oisillons selon la taille de la couvée et leurs intervalles de confiance à 95%. Les triangles représentent les effets estimés par les différentes équipes de recherche et le cercle indique l'effet moyen de l'ensemble des équipes. On constate un consensus d'un effet négatif de la taille de la couvée sur la croissance des oisillons vers  $-0.35 \pm 0.06$  et une grande variation entre -1.55 et 0.38. Tiré de Gould et al., 2023

## *Discussion générale      Transmission de l'incertitude scientifique face à la prise d'action*

Bien que cette incertitude puisse freiner la prise de décision<sup>42</sup>, elle n'en est pas le seul obstacle (Maxim & Van Der Sluijs, 2007). L'action de conservation peut aussi être entravée par principe de précaution, qui part du postulat que le contexte social favorise l'exploitation d'une ressource ou d'un milieu pour favoriser le développement, *i.e.* le bien-être humain, et qu'une action opposée est perçue comme un obstacle à ce développement.

*'Under this presumption, where there is uncertainty or ignorance regarding the impacts of an activity such as release of pollutants, fishing, building, or mining, the "default state" is that activities can go ahead. Uncertainty around environmental impact is used as a rationale for not banning toxic chemicals, not reducing fisheries harvest levels, or not refusing mining applications. Environmental objections against them will require clear scientific evidence that they lead to environmental harm.'* – Cooney, 2005, 2. The meaning of the precautionary principle, In 'Biodiversity and the Precautionary Principle: Risk and Uncertainty in Conservation and Sustainable Use'.

Afin d'illustrer ce propos, dans le contexte de l'avifaune française, on pourra prendre, par exemple, la nécessité de renouveler annuellement les arrêtés d'interdiction de chasse (LPO France, 2024) pour le Courlis cendré (*Numenius arquata*) et la Barge à queue noire (*Limosa limosa*), malgré, respectivement, le constat d'une population en déclin depuis 30 ans (Viana et al., 2023) et des pratiques de conservation à l'échelle européenne jugées inadéquates (van der Vliet et al., 2015). Ou encore, les requêtes de dérogations des chasseurs afin de légaliser la chasse du Bruant ortolan (*Emberiza hortulana*), une espèce en danger d'extinction (EN ; UICN France et al., 2016) dans de multiples pays européens et en France (Jiguet et al., 2016), à des fins traditionnelles (Jiguet et al., 2019).

Les sciences de la conservation se retrouvent dans un contexte de tension mêlant (i) l'étude de dynamique et la production d'indicateurs n'évoluant pas sur les mêmes pas-de-

---

<sup>42</sup> Illustré par la citation de l'article de presse autour sur le cas des néonicotinoïde (Bodoni, 2017).

## *Discussion générale      Transmission de l'incertitude scientifique face à la prise d'action*

temps que le court-termisme de la prise de décision (Levrel, 2007) ; et (ii) le manque de financement face à l'ampleur de la crise de la biodiversité (Deutz et al., 2020; Herweijer et al., 2020). Cette tension voit émerger des remises en question de l'allocation de la ressource disponible entre la production scientifique et le financement de mesures de conservation (McDonald-Madden et al., 2010; Buxton et al., 2020). Le recours aux sciences citoyennes, en association avec des modèles intégrés (Robinson et al., 2018; Sun, Royle & Fuller, 2019) semble offrir des perspectives prometteuses pour réagir face à ce contexte (Buxton et al., 2021; Castagneyrol, Bedessem & Julliard, 2023) tout en sensibilisant le public aux causes environnementales (Bonney et al., 2009).

## VII Références de l'introduction générale

---

- Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:716–723. DOI: 10.1109/TAC.1974.1100705.
- Aldrich J. 1997. R.A. Fisher and the making of maximum likelihood 1912-1922. *Statistical Science* 12:162–176. DOI: 10.1214/ss/1030037906.
- Almond REA, Grooten M, Juffe B, Peterson T. 2022. Living Planet Report 2022 – Building a naturepositive society. WWF.
- Alsila T, Elo M, Hakkari T, Kotiaho JS. 2021. Effects of habitat restoration on peatland bird communities. *Restoration Ecology* 29:e13304. DOI: 10.1111/rec.13304.
- Aubry P, Francesiaz C. 2022. On comparing design-based estimation versus model-based prediction to assess the abundance of biological populations. *Ecological Indicators* 144:109394. DOI: 10.1016/j.ecolind.2022.109394.
- Aubry P, Quaintenne G, Dupuy J, Francesiaz C, Guillemain M, Caizergues A. 2023. On using stratified two-stage sampling for large-scale multispecies surveys. *Ecological Informatics* 77:102229. DOI: 10.1016/j.ecoinf.2023.102229.
- Audubon JJ. 1831. Passenger pigeon. In: *Ornithological biography*. Edinburgh: Adam Black., DOI: 10.5962/bhl.title.48976.
- Bailey FM. 1898. *Birds through an opera-glass*. Boston: Houghton, Mifflin. DOI: 10.5962/bhl.title.13728.
- Barker RJ, Schofield MR, Link WA, Sauer JR. 2018. On the Reliability of N-mixture Models for Count Data. *Biometrics* 74:369–377. DOI: 10.1111/biom.12734.
- Barnagaud J-Y, Devictor V, Jiguet F, Barbet-Massin M, Viol IL, Archaux F. 2012. Relating Habitat and Climatic Niches in Birds. *PLOS ONE* 7:e32819. DOI: 10.1371/journal.pone.0032819.
- Bayraktarov E, Ehmke G, O'Connor J, Burns EL, Nguyen HA, McRae L, Possingham HP, Lindenmayer DB. 2019. Do Big Unstructured Biodiversity Data Mean More Knowledge? *Frontiers in Ecology and Evolution* 6.
- Berthold P. 1999. A comprehensive theory for the evolution, control and adaptability of avian migration. *Ostrich* 70:1–11. DOI: 10.1080/00306525.1999.9639744.
- Blanchet FG, Cazelles K, Gravel D. 2020. Co-occurrence is not evidence of ecological interactions. *Ecology Letters* 23:1050–1063. DOI: 10.1111/ele.13525.
- Blockstein DE, Tordoff B. 1985. Was overhunting the cause of the Passenger Pigeon's demise or was the bird an early victim of technology?
- Blondel J, Ferry C, Frochot B. 1981. *Point Counts with Unlimited Distance*. Studies in Avian Biology.
- Boakes EH, McGowan PJK, Fuller RA, Chang-qing D, Clark NE, O'Connor K, Mace GM. 2010. Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data. *PLOS Biology* 8:e1000385. DOI: 10.1371/journal.pbio.1000385.
- Bodey TW, Cleasby IR, Bell F, Parr N, Schultz A, Votier SC, Bearhop S. 2018. A phylogenetically controlled meta-analysis of biologging device effects on birds:

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Deleterious effects and a call for more standardized reporting of study data. *Methods in Ecology and Evolution* 9:946–955. DOI: 10.1111/2041-210X.12934.
- Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips T, Rosenberg KV, Shirk J. 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience* 59:977–984. DOI: 10.1525/bio.2009.59.11.9.
- Borchers D, Buckland S, Zucchini W. 2002a. 3. Building blocks. In: *Estimating Animal Abundance: Closed Populations*. DOI: 10.1007/978-1-4471-3708-5.
- Borchers D, Buckland S, Zucchini W. 2002b. 2. Using likelihood for estimation. In: *Estimating Animal Abundance: Closed Populations*. DOI: 10.1007/978-1-4471-3708-5.
- Bottrill MC, Joseph LN, Carwardine J, Bode M, Cook C, Game ET, Grantham H, Kark S, Linke S, McDonald-Madden E, Pressey RL, Walker S, Wilson KA, Possingham HP. 2008. Is conservation triage just smart decision making? *Trends in Ecology & Evolution* 23:649–654. DOI: 10.1016/j.tree.2008.07.007.
- Bottrill MC, Joseph LN, Carwardine J, Bode M, Cook C, Game ET, Grantham H, Kark S, Linke S, McDonald-Madden E, Pressey RL, Walker S, Wilson KA, Possingham HP. 2009. Finite conservation funds mean triage is unavoidable. *Trends in Ecology & Evolution* 24:183–184. DOI: 10.1016/j.tree.2008.11.007.
- Bowerman WW, Giesy JP, Best DA, Kramer VJ. 1995. A review of factors affecting productivity of bald eagles in the Great Lakes region: implications for recovery. *Environmental Health Perspectives* 103:51–59.
- Bowler DE, Bjorkman AD, Dornelas M, Myers-Smith IH, Navarro LM, Niamir A, Supp SR, Waldo C, Winter M, Vellend M, Blowes SA, Böhning-Gaese K, Bruelheide H, Elahi R, Antão LH, Hines J, Isbell F, Jones HP, Magurran AE, Cabral JS, Bates AE. 2020. Mapping human pressures on biodiversity across the planet uncovers anthropogenic threat complexes. *People and Nature* 2:380–394. DOI: 10.1002/pan3.10071.
- Bravington MV, Miller DL, Hedley SL. 2021. Variance propagation for density surface models. *Journal of Agricultural, Biological and Environmental Statistics* 26:306–323. DOI: 10.1007/s13253-021-00438-2.
- Bray JR, Curtis JT. 1957. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* 27:325–349. DOI: 10.2307/1942268.
- Bromaghin JF, McDonald TL, Amstrup SC. 2013. Plausible combinations: An improved method to evaluate the covariate structure of Cormack-Jolly-Seber mark-recapture models. *Open Journal of Ecology* 3:11–22. DOI: 10.4236/oje.2013.31002.
- Bucher EH. 1992. The Causes of Extinction of the Passenger Pigeon. In: Power DM ed. *Current Ornithology*. Boston, MA: Springer US, 1–36. DOI: 10.1007/978-1-4757-9921-7\_1.
- Buckland ST, Anderson DR, Burnham KP, Laake JL. 1993a. 3. Statistical theory. In: *Distance Sampling*. Dordrecht: Springer Netherlands,. DOI: 10.1007/978-94-011-1574-2.
- Buckland ST, Anderson DR, Burnham KP, Laake JL. 1993b. 1. Introductory concepts. In: *Distance Sampling*. Dordrecht: Springer Netherlands,. DOI: 10.1007/978-94-011-1574-2.
- Buckland S, Anderson D, Burnham K, Laake J, Borchers D, Thomas L. 2004. *Advanced Distance Sampling*. DOI: 10.1093/oso/9780198507833.001.0001.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Buckland ST, Johnston A. 2017. Monitoring the biodiversity of regions: Key principles and possible pitfalls. *Biological Conservation* 214:23–34. DOI: 10.1016/j.biocon.2017.07.034.
- Burnham KP. 1981. SUMMARIZING REMARKS: ENVIRONMENTAL.
- Burnham KP, Anderson DR. 2001. Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research* 28:111–119. DOI: 10.1071/wr99107.
- Burnham KP, Anderson DR, Huyvaert KP. 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* 65:23–35. DOI: 10.1007/s00265-010-1029-6.
- Butcher GS, Robbins CS, Bystrak D, Geissler PH. 1987. The Breeding Bird Survey: Its First Fifteen Years, 1965-1979. *The Condor* 89:681. DOI: 10.2307/1368666.
- Buzhdygan O, Patten B, Rudenko S. 2012. 12. Trophic Network Analysis: Comparison of System-Wide Properties. In: Jordan F, Jørgensen SE eds. *Models of the Ecological Hierarchy: From Molecules to the Ecosphere*. Amsterdam ; Boston: Elsevier,.
- Calabrese JM, Certain G, Kraan C, Dormann CF. 2014. Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography* 23:99–112. DOI: <https://doi.org/10.1111/geb.12102>.
- Callaghan CT, Bowler DE, Blowes SA, Chase JM, Lyons MB, Pereira HM. 2022. Quantifying effort needed to estimate species diversity from citizen science data. *Ecosphere* 13:e3966. DOI: 10.1002/ecs2.3966.
- Carraro L, Hartikainen H, Jokela J, Bertuzzo E, Rinaldo A. 2018. Estimating species distribution and abundance in river networks using environmental DNA. *Proceedings of the National Academy of Sciences* 115:11724–11729. DOI: 10.1073/pnas.1813843115.
- Casella G, Berger RL. 2001. 7. Point estimation. In: *Statistical Inference*. Pacific Grove, Calif: Cengage Learning,.
- Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM. 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances* 1:e1400253. DOI: 10.1126/sciadv.1400253.
- Challear M, Lavarde P. 2014. Les plans nationaux d'actions en faveur des espèces menacées : Une politique à refonder.
- Chamberlin TC. 1890. The Method of Multiple Working Hypotheses. *Science* 15:92–96.
- Chambert T, Miller DAW, Nichols JD. 2015. Modeling false positive detections in species occurrence data under different study designs. *Ecology* 96:332–339. DOI: 10.1890/14-1507.1.
- Chandler RB, Royle JA, King DI. 2011. Inference about density and temporary emigration in unmarked populations. *Ecology* 92:1429–1435. DOI: 10.1890/10-2433.1.
- Chapman FM (ed.). 1901. *Bird Lore*. Harrisburg, New York city.
- Clamens A, LPO Auvergne. 2010. *Atlas des oiseaux nicheurs d'Auvergne*. Paris: DELACHAUX.
- Coelho MTP, Diniz-Filho JA, Rangel TF. 2019. A parsimonious view of the parsimony principle in ecology and evolution. *Ecography* 42:968–976. DOI: 10.1111/ecog.04228.
- Collen B, Loh J, Whitmee S, McRAE L, Amin R, Baillie JEM. 2009. Monitoring Change in Vertebrate Abundance: the Living Planet Index. *Conservation Biology* 23:317–327. DOI: 10.1111/j.1523-1739.2008.01117.x.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Colwell RK, Rangel TF. 2009. Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences* 106:19651–19658. DOI: 10.1073/pnas.0901650106.
- Comolet-Tirman J, Quaintenne G, Sibley J-P, Wroza S, Bal G, Champagnon J, Couzi L, Czajkowski M-A, Denis P, Frochot B, Muller Y. 2022. Estimation des populations d'oiseaux nicheurs de France métropolitaine. *Alauda* 90:133–150.
- Comolet-Tirman J, Sibley J-P, Witté I, Cadiou B, Czajkowski M-A, Deceuninck B, Jiguet F, Landry P, Quaintenne G, Roché J-E, Sarasa M, Touroult J. 2015. STATUTS ET TENDANCES DES POPULATIONS D'OISEAUX NICHEURS DE FRANCE Bilan simplifié du premier rapportage national au titre de la Directive Oiseaux. 83.
- Conn PB, Johnson DS, Boveng PL. 2015. On Extrapolating Past the Range of Observed Data When Making Statistical Predictions in Ecology. *PLOS ONE* 10:e0141416. DOI: 10.1371/journal.pone.0141416.
- Crick HQP. 1992. A bird-habitat coding system for use in Britain and Ireland incorporating aspects of land-management and human activity. *Bird Study* 39:1–12. DOI: 10.1080/00063659209477092.
- Crowley MA, Cardille JA. 2020. Remote Sensing's Recent and Future Contributions to Landscape Ecology. *Current Landscape Ecology Reports* 5:45–57. DOI: 10.1007/s40823-020-00054-9.
- Crutzen PJ, Stoermer EF. 2000. The “Anthropocene.”
- Davidson NC, Stroud DA. 2006. African-Western Eurasian Flyways: current knowledge, population status and future challenges. In: Boere GC, Galbraith CA, Stroud DA eds. *Waterbirds around the world*. Edinburgh, UK, 63–73.
- Delisle ZJ, Flaherty EA, Nobbe MR, Wzientek CM, Swihart RK. 2021. Next-Generation Camera Trapping: Systematic Review of Historic Trends Suggests Keys to Expanded Research Applications in Ecology and Conservation. *Frontiers in Ecology and Evolution* 9. DOI: 10.3389/fevo.2021.617996.
- Deplazes-Zemp A. 2024. Beyond Intrinsic and Instrumental: Third-Category Value in Environmental Ethics and Environmental Policy. *Ethics, Policy & Environment* 27:166–188. DOI: 10.1080/21550085.2023.2166341.
- Deutz A, Heal GM, Niu R, Swanson E, Townsend T, Zhu L, Delmar A, Meghji A, Sethi SA, Tobin de la Puente J. 2020. *Financing Nature: Closing the global biodiversity financing gap*. The Paulson Institute, The Nature Conservancy, and the Cornell Atkinson Center for Sustainability.
- Devictor V, Whittaker RJ, Beltrame C. 2010. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions* 16:354–362. DOI: 10.1111/j.1472-4642.2009.00615.x.
- Directive 2009/147/CE du Parlement européen et du Conseil du 30 novembre 2009 concernant la conservation des oiseaux sauvages (Version codifiée). 2009.
- Dumelle M, Higham M, Ver Hoef JM, Olsen AR, Madsen L. 2022. A comparison of design-based and model-based approaches for finite population spatial sampling and inference. *Methods in Ecology and Evolution* 13:2018–2029. DOI: 10.1111/2041-210X.13919.
- Dunn AM, Weston MA. 2008. A review of terrestrial bird atlases of the world and their application. *Emu - Austral Ornithology* 108:42–67. DOI: 10.1071/MU07034.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Dupuis-Tate M-F, Fischesser B. 2017. *Le Guide illustré de l'écologie*. Paris: DELACHAUX.
- Efron B. 2012. A 250-year argument: Belief, behavior, and the bootstrap. *Bulletin of the American Mathematical Society* 50:129–146. DOI: 10.1090/S0273-0979-2012-01374-5.
- Ehrlich PR, Ehrlich AH. 1992. The Value of Biodiversity. *Ambio* 21:219–226.
- Ehrlich PR, Mooney HA. 1983. Extinction, Substitution, and Ecosystem Services. *BioScience* 33:248–254. DOI: 10.2307/1309037.
- Elith J, Leathwick JR. 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics* 40:677–697. DOI: 10.1146/annurev.ecolsys.110308.120159.
- Ellison AM. 2004. Bayesian inference in ecology. *Ecology Letters* 7:509–520. DOI: 10.1111/j.1461-0248.2004.00603.x.
- Elton C. 1927. The animal community. In: *Animal ecology*.
- Emlen JT. 1971. Population Densities of Birds Derived from Transect Counts. *The Auk* 88:323–342. DOI: 10.2307/4083883.
- Enemar A, Sjöstrand B, Andersson G, Proschwitz T von. 2004. The 37-year dynamics of a subalpine passerine bird community, with special emphasis on the influence of environmental temperature and *Epirrita autumnata* cycles. *Ornis Svecica* 14:63–106. DOI: 10.34080/os.v14.20236.
- Farmer RG, Leonard ML, Horn AG. 2012. Observer Effects and Avian-Call-Count Survey Quality: Rare-Species Biases and Overconfidence. *The Auk* 129:76–86. DOI: 10.1525/auk.2012.11129.
- Farmer RG, Leonard ML, Mills Flemming JE, Anderson SC. 2014. Observer aging and long-term avian survey data quality. *Ecology and Evolution* 4:2563–2576. DOI: 10.1002/ece3.1101.
- Fath B, Patten B. 1999. Review of the Foundations of Network Environ Analysis. *Ecosystems* 2:167–179. DOI: 10.1007/s100219900067.
- Fiedler W. 2009. Bird Ecology as an Indicator of Climate and Global Change. In: *Climate Change*. Elsevier, 181–195. DOI: 10.1016/B978-0-444-53301-2.00009-9.
- Fitter M, Fitter RSR (eds.). 1987. *The road to extinction : problems of categorizing the status of taxa threatened with extinction : proceedings of a symposium held by the Species Survival Commission, Madrid, 7 and 9 November 1984*. IUCN.
- Fitzpatrick MC, Preisser EL, Ellison AM, Elkinton JS. 2009. Observer bias and the detection of low-density populations. *Ecological Applications* 19:1673–1679. DOI: 10.1890/09-0265.1.
- Fletcher RJ, Hefley TJ, Robertson EP, Zuckerberg B, McCleery RA, Dorazio RM. 2019. A practical guide for combining data to model species distributions. *Ecology* 100:e02710. DOI: 10.1002/ecy.2710.
- Fontaine B, Moussy C, Carricaburu JC, Dupuis J, Schmaltz L, Lorrillière R, Loïs G, Gaudard C. 2020. *Suivi des oiseaux communs en France 1989-2019 : 30 ans de suivis participatifs*.
- Fuller RJ, Langslow DR. 1984. Estimating numbers of birds by point counts: how long should counts last? *Bird Study* 31:195–202. DOI: 10.1080/00063658409476841.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Gaston KJ, Cox DTC, Canavelli SB, García D, Hughes B, Maas B, Martínez D, Ogada D, Inger R. 2018. Population Abundance and Ecosystem Service Provision: The Case of Birds. *BioScience* 68:264–272. DOI: 10.1093/biosci/biy005.
- Gaston KJ, Fuller RA. 2008. Commonness, population depletion and conservation biology. *Trends in Ecology & Evolution* 23:14–19. DOI: 10.1016/j.tree.2007.11.001.
- Gauthreaux SA Jr, Belser CG. 2003. Radar Ornithology and Biological Conservation. *The Auk* 120:266–277. DOI: 10.1093/auk/120.2.266.
- Gaüzère P, Devictor V. 2021. Mismatches between birds' spatial and temporal dynamics reflect their delayed response to global changes. *Oikos* 130:1284–1296. DOI: 10.1111/oik.08289.
- Gaüzère P, Doucier G, Devictor V, Kéfi S. 2019. A framework for estimating species-specific contributions to community indicators. *Ecological Indicators* 99:74–82. DOI: 10.1016/j.ecolind.2018.11.069.
- Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. 2013a. 1. Probability and inference. In: *Bayesian Data Analysis*. Chapman and Hall/CRC,. DOI: 10.1201/b16018.
- Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. 2013b. 11. Basics of Markov chain simulation. In: *Bayesian Data Analysis*. Chapman and Hall/CRC,. DOI: 10.1201/b16018.
- Gelman A, Rubin DB. 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* 7:457–472. DOI: 10.1214/ss/1177011136.
- Gibb R, Browning E, Glover-Kapfer P, Jones KE. 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution* 10:169–185. DOI: 10.1111/2041-210X.13101.
- Gibbons DW, Donald PF, Bauer H-G, Fornasari L, Dawson IK. 2007. Mapping avian distributions: the evolution of bird atlases. *Bird Study* 54:324–334. DOI: 10.1080/00063650709461492.
- Gilks WR, Richardson S, Spiegelhalter D (eds.). 1995. *Markov Chain Monte Carlo in Practice*. New York: Chapman and Hall/CRC. DOI: 10.1201/b14835.
- Gilpin ME, Soulé ME. 1986. 2. Minimum viable populations: Processes of species extinction. In: Soulé ME ed. *Conservation biology: the science of scarcity and diversity*. Sunderland, Mass: Sinauer Associates, 19–34.
- Godet L, Jaffré M, Devictor V. 2011. Waders in winter: long-term changes of migratory bird assemblages facing climate change. *Biology Letters* 7:714–717. DOI: 10.1098/rsbl.2011.0152.
- Gotelli NJ, Ellison AM. 2012. 4. Framing and Testing Hypotheses. In: *A Primer of Ecological Statistics*. Sunderland, Mass: Sinauer Associates is an imprint of Oxford University Press,.
- Gregory RD, Bashford RI, Balmer DE, Marchant JH, Wilson AM, Baillie SR. 1996. *The Breeding Bird Survey 1994-1995*. Thetford: British Trust for Ornithology.
- Grinnell J, Storer TI. 1924. Distribution of animal life in the Yosemite section. In: *Animal life in the Yosemite; an account of the mammals, birds, reptiles, and amphibians in a cross-section of the Sierra Nevada*. Berkeley, Calif: University of California Press,. DOI: 10.5962/bhl.title.32895.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Guillera-Arroita G. 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography* 40:281–295. DOI: 10.1111/ecog.02445.
- Gurevitch J, Padilla DK. 2004. Are invasive species a major cause of extinctions? *Trends in Ecology & Evolution* 19:470–474. DOI: 10.1016/j.tree.2004.07.005.
- Halliday TR. 1980. The extinction of the passenger pigeon *ectopistes migratorius* and its relevance to contemporary conservation. *Biological Conservation* 17:157–162. DOI: 10.1016/0006-3207(80)90046-4.
- Hampton SE, Strasser CA, Tewksbury JJ, Gram WK, Budden AE, Batcheller AL, Duke CS, Porter JH. 2013. Big data and the future of ecology. *Frontiers in Ecology and the Environment* 11:156–162. DOI: 10.1890/120103.
- Hastie T, Tibshirani R, Friedman J. 2009a. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. New York, NY: Springer.
- Hastie T, Tibshirani R, Friedman J. 2009b. 7. Model Assessment and Selection. In: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. New York, NY: Springer,.
- Hedley SL, Buckland ST. 2004. Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 9:181–199. DOI: 10.1198/1085711043578.
- Hellou J, Lebeuf M, Rudi M. 2013. Review on DDT and metabolites in birds and mammals of aquatic ecosystems. *Environmental Reviews* 21:53–69. DOI: 10.1139/er-2012-0054.
- Herweijer C, Evison W, Mariam S, Khatri A, Albani M, Semov A, Long E. 2020. *Nature Risk Rising: Why the Crisis Engulfing Nature Matters for Business and the Economy*.
- Heus M. 2013. Migration in birds: Why go, where to and how to get there.
- Hilborn R, Mangel M. 1997. 1. An Ecological Scenario and the Tools of the Ecological Detective. In: *The Ecological Detective: Confronting Models with Data*. Princeton, NJ: Princeton University Press,.
- Himes A, Muraca B. 2018. Relational values: the key to pluralistic valuation of ecosystem services. *Current Opinion in Environmental Sustainability*.
- Hirzel AH, Le Lay G. 2008. Habitat suitability modelling and niche theory. *Journal of Applied Ecology* 45:1372–1381. DOI: 10.1111/j.1365-2664.2008.01524.x.
- Hobbs NT, Hooten M. 2015. 6. Hierarchical Bayesian Models. In: *Bayesian Models: A Statistical Primer for Ecologists*. Princeton, New Jersey: Princeton University Press,.
- Hoekstra JM, Boucher TM, Ricketts TH, Roberts C. 2005. Confronting a biome crisis: global disparities of habitat loss and protection. *Ecology Letters* 8:23–29. DOI: 10.1111/j.1461-0248.2004.00686.x.
- Hooper DU, Adair EC, Cardinale BJ, Byrnes JEK, Hungate BA, Matulich KL, Gonzalez A, Duffy JE, Gamfeldt L, O'Connor MI. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* 486:105–108. DOI: 10.1038/nature11118.
- Hornaday WT. 1913a. II. Extinct species of north american birds. In: *Our Vanishing Wildlife: Its Extermination And Preservation*. Kessinger Publishing,.
- Hornaday WT. 1913b. XXIII. The economic value of birds. In: *Our Vanishing Wildlife: Its Extermination And Preservation*. Kessinger Publishing, 213–233.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Howes C, Symes CT, Byholm P. 2019. Evidence of large-scale range shift in the distribution of a Palaearctic migrant in Africa. *Diversity and Distributions* 25:1142–1155. DOI: 10.1111/ddi.12922.
- Hutchinson GE. 1957. Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22:415–427. DOI: 10.1101/SQB.1957.022.01.039.
- Iknayan KJ, Tingley MW, Furnas BJ, Beissinger SR. 2014. Detecting diversity: emerging methods to estimate species diversity. *Trends in Ecology & Evolution* 29:97–106. DOI: 10.1016/j.tree.2013.10.012.
- Issa N, Muller Y. 2015. Atlas des oiseaux de France métropolitaine: Nidification et présence hivernale. Paris: DELACHAUX.
- IUCN. 2012. Catégories et Critères de la Liste rouge. IUCN.
- Jarzyna MA, Jetz W. 2016. Detecting the Multiple Facets of Biodiversity. *Trends in Ecology & Evolution* 31:527–538. DOI: 10.1016/j.tree.2016.04.002.
- Jeliazkov A, Gavish Y, Marsh CJ, Geschke J, Brummitt N, Rocchini D, Haase P, Kunin WE, Henle K. 2022. Sampling and modelling rare species: Conceptual guidelines for the neglected majority. *Global Change Biology* 28:3754–3777. DOI: 10.1111/gcb.16114.
- Jetz W, McGeoch MA, Guralnick RP, Ferrier S, Beck J, Costello MJ, Fernandez M, Miguel Fernandez, Miguel Fernandez, Fernández M, Geller GN, Keil P, Merow C, Meyer C, Muller-Karger FE, Pereira HM, Regan EC, Regan E, Schmeller DS, Turak E. 2019. Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology and Evolution* 3:539–551. DOI: 10.1038/s41559-019-0826-1.
- Jiguet F, Devictor V, Julliard R, Couvet D. 2012. French citizens monitoring ordinary birds provide tools for conservation and ecological sciences. *Acta Oecologica* 44:58–66. DOI: 10.1016/j.actao.2011.05.003.
- Jiguet F, Devictor V, Ottvall R, Turnhout CV, Jeugd HV der, Lindström Å. 2010. Bird population trends are linearly affected by climate change along species thermal ranges. *Proceedings of the Royal Society B: Biological Sciences*. DOI: 10.1098/rspb.2010.0796.
- Johnson CN, Balmford A, Brook BW, Buettel JC, Galetti M, Guangchun L, Wilmshurst JM. 2017. Biodiversity losses and conservation responses in the Anthropocene. *Science* 356:270–275. DOI: 10.1126/science.aam9317.
- Johnson JB, Omland KS. 2004. Model selection in ecology and evolution. *Trends in Ecology & Evolution* 19:101–108. DOI: 10.1016/j.tree.2003.10.013.
- Johnston A, Hochachka WM, Strimas-Mackey ME, Gutierrez VR, Robinson OJ, Miller ET, Auer T, Kelling ST, Fink D. 2019. Best practices for making reliable inferences from citizen science data: case study using eBird to estimate species distributions. *bioRxiv:574392*. DOI: 10.1101/574392.
- Johnston A, Hochachka WM, Strimas-Mackey ME, Gutierrez VR, Robinson OJ, Miller ET, Auer T, Kelling ST, Fink D. 2020. Analytical guidelines to increase the value of citizen science data: using eBird data to estimate species occurrence. *bioRxiv:574392*. DOI: 10.1101/574392.
- Johnston A, Matechou E, Dennis EB. 2022. Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution* 14:103–116. DOI: 10.1111/2041-210X.13834.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Johnston A, Newson SE, Risely K, Musgrove AJ, Massimino D, Baillie SR, Pearce-Higgins JW. 2014. Species traits explain variation in detectability of UK birds. *Bird Study* 61:340–350. DOI: 10.1080/00063657.2014.941787.
- Jolliffe IT, Cadima J. 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374:20150202. DOI: 10.1098/rsta.2015.0202.
- Joly D, Brossard T, Cardot H, Cavailles J, Hilal M, Wavresky P. 2010. Les types de climats en France, une construction spatiale. *Cybergeo: European Journal of Geography*. DOI: 10.4000/cybergeo.23155.
- Josselyn J. 1674. *An Account of Two Voyages to New-England, Made during the Years 1638, 1663*. Wisconsin historical society.
- Jourde P. 2015. Biogéographie de la France métropolitaine. In: *Atlas des oiseaux de France métropolitaine: Nidification et présence hivernale*. Paris: DELACHAUX, 23–29.
- Julliard R, Jiguet F. 2002. Un suivi intégré des populations d'oiseaux communs en France. *Alauda*:137–147.
- Junge MRJ, Dettori JR. 2018. ROC Solid: Receiver Operator Characteristic (ROC) Curves as a Foundation for Better Diagnostic Tests. *Global Spine Journal* 8:424–429. DOI: 10.1177/2192568218778294.
- Keller V, Herrando S, Voříšek P, Franch M, Kipson M, Milanese P, Martí D, Anton M, Klvaňová A, Kalyakin MV, Bauer H-G, Foppen RPB. 2020. *European Breeding Bird Atlas 2: Distribution, Abundance and Change*. European Bird Census Council & Lynx Edicions, Barcelona.
- Kelling S, Fink D, Hochachka W, Iliff M, Sullivan B, Wood C, Munson A, Riedewald M. 2009. Obtaining New Insights for Biodiversity Conservation from Broad-Scale Citizen Science Data. *Nature Precedings*:1–1. DOI: 10.1038/npre.2009.3967.1.
- Kelling S, Johnston A, Bonn A, Fink D, Ruiz-Gutierrez V, Bonney R, Fernandez M, Hochachka WM, Julliard R, Kraemer R, Guralnick R. 2019. Using Semistructured Surveys to Improve Citizen Science Data for Monitoring Biodiversity. *BioScience* 69:170–179. DOI: 10.1093/biosci/biz010.
- Kellner KF, Swihart RK. 2014. Accounting for Imperfect Detection in Ecology: A Quantitative Review. *PLOS ONE* 9:e111436. DOI: 10.1371/journal.pone.0111436.
- Kendall WL, Peterjohn BG, Sauer JR. 1996. First-Time Observer Effects in the North American Breeding Bird Survey. *The Auk* 113:823–829. DOI: 10.2307/4088860.
- Kéry M, Royle JA. 2015a. 1. Distribution, Abundance and Species Richness in Ecology. In: *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1:Prelude and Static Models*. Amsterdam ; Boston: Academic Press,.
- Kéry M, Royle JA. 2015b. 2. What are Hierarchical models and how do we analyze them? In: *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1:Prelude and Static Models*. Amsterdam ; Boston: Academic Press,.
- Kéry M, Royle JA. 2015c. 8. Modeling abundance using Hierarchical Distance Sampling. In: *Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and*

- species richness in R and BUGS: Volume 1:Prelude and Static Models. Amsterdam ; Boston: Academic Press,.
- Kéry M, Royle JA. 2015d. 9.5. Open HDS models: Temporary emigration. In: Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1:Prelude and Static Models. Amsterdam ; Boston: Academic Press,.
- Kéry M, Royle JA. 2015e. 9. Advanced Hierarchical Distance Sampling. In: Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1:Prelude and Static Models. Amsterdam ; Boston: Academic Press,.
- Kéry M, Royle JA. 2015f. 5.10 GoF Assessment: Posterior Predictive Checks and the Parametric Bootstrap. In: Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1:Prelude and Static Models. Amsterdam ; Boston: Academic Press,.
- Kéry M, Royle JA, Hallman T, Robinson WD, Strebel N, Kellner KF. 2024. Integrated distance sampling models for simple point counts. *Ecology* n/a:e4292. DOI: 10.1002/ecy.4292.
- Kéry M, Schaub M. 2011. Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective. Boston: Academic Press.
- King FW. 1978. The wildlife trade. In: Brokaw HP ed. *Wildlife and America: Contributions to an Understanding of American Wildlife and Its Conservation*. Council on Environmental Quality, 253–271.
- King R. 2014. Statistical Ecology. *Annual Review of Statistics and Its Application* 1:401–426. DOI: 10.1146/annurev-statistics-022513-115633.
- Kullback S, Leibler RA. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics* 22:79–86. DOI: 10.1214/aoms/1177729694.
- La Morgia V, Calmanti R, Calabrese A, Focardi S. 2015. Cost-effective nocturnal distance sampling for landscape monitoring of ungulate populations. *European Journal of Wildlife Research* 61:285–298. DOI: 10.1007/s10344-014-0898-9.
- La Sorte FA, Lepczyk CA, Burnett JL, Hurlbert AH, Tingley MW, Zuckerberg B. 2018. Opportunities and challenges for big data ornithology. *The Condor* 120:414–426. DOI: 10.1650/CONDOR-17-206.1.
- Ladle RJ, Jepson P, Gilson L. 2011. 2. Social values and Conservation biogeography. In: *Conservation Biogeography*. John Wiley & Sons,.
- Lefort M-C, Cruickshank RH, Descovich K, Adams NJ, Barun A, Emami-Khoyi A, Ridden J, Smith VR, Sprague R, Waterhouse B, Boyer S. 2022. Blood, sweat and tears: a review of non-invasive DNA sampling. *Peer Community Journal* 2. DOI: 10.24072/pcjournal.98.
- Lehikoinen A, Lindén A, Karlsson M, Andersson A, Crewe TL, Dunn EH, Gregory G, Karlsson L, Kristiansen V, Mackenzie S, Newman S, Røer JE, Sharpe C, Sokolov LV, Steinholtz Å, Stervander M, Tirri I-S, Tjørnløv RS. 2019. Phenology of the avian spring migratory passage in Europe and North America: Asymmetric advancement in time and increase in duration. *Ecological Indicators* 101:985–991. DOI: 10.1016/j.ecolind.2019.01.083.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Shurin JB, Law R, Tilman D, Loreau M, Gonzalez A. 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* 7:601–613. DOI: 10.1111/j.1461-0248.2004.00608.x.
- Levin SA. 1992. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology* 73:1943–1967. DOI: 10.2307/1941447.
- Loh J, Green RE, Ricketts T, Lamoreux J, Jenkins M, Kapos V, Randers J. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360:289–295. DOI: 10.1098/rstb.2004.1584.
- Luque S, Pettorelli N, Vihervaara P, Wegmann M. 2018. Improving biodiversity monitoring using satellite remote sensing to provide solutions towards the 2020 conservation targets. *Methods in Ecology and Evolution* 9:1784–1786. DOI: 10.1111/2041-210X.13057.
- Mace GM. 2011. Drivers of Biodiversity Change. In: Leader-Williams N, Adams WM, Smith RJ eds. *Trade-offs in Conservation: Deciding What to Save*. John Wiley & Sons,.
- Mace GM, Possingham HP, Leader-Williams N. 2008. 2. Prioritizing choices in conservation. In: Macdonald DW ed. *Key topics in conservation biology*. Malden, MA: Blackwell,.
- Magurran AE, Baillie SR, Buckland ST, Dick JMcP, Elston DA, Scott EM, Smith RI, Somerfield PJ, Watt AD. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution* 25:574–582. DOI: 10.1016/j.tree.2010.06.016.
- Marcel C, Guéguen J, Body G, Villiers A. 2021. Dataviz - Observer les oiseaux nicheurs et hivernants présents en métropole.
- Marion L, Vessem J, Ulenaers P. 2000. Herons in Europe. In: 1–31.
- Marques FFC, Buckland ST. 2003. Incorporating Covariates into Standard Line Transect Analyses. *Biometrics* 59:924–935. DOI: 10.1111/j.0006-341X.2003.00107.x.
- Martin ET. 1879. The Pigeon Butcher's Defense. In: *The Passenger Pigeon*.
- Matutini F, Baudry J, Pain G, Sineau M, Python J. 2021. How citizen science could improve species distribution models and their independent assessment. *Ecology and Evolution* 11:3028–3039. DOI: 10.1002/ece3.7210.
- Maurer BA, McGill BJ. 2010. Measurement of species diversity. In: Magurran AE, McGill BJ eds. *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford, New York: Oxford University Press, 55–64.
- Mayaud N, Balsac HH de, Jouard H. 1936. *Inventaire des oiseaux de France*. Paris, France: Société d'études ornithologiques.
- McClintock BT, Thomas L. 2020. 3. Estimating Abundance or Occupancy from Unmarked Populations. In: Murray DL, Sandercock BK eds. *Population Ecology in Practice*. Hoboken: Wiley-Blackwell,.
- McKinley DC, Miller-Rushing AJ, Ballard HL, Bonney R, Brown H, Cook-Patton SC, Evans DM, French RA, Parrish JK, Phillips TB, Ryan SF, Shanley LA, Shirk JL, Stepenuck KF, Weltzin JF, Wiggins A, Boyle OD, Briggs RD, Chapin SF, Hewitt DA, Preuss PW, Soukup MA. 2017. Citizen science can improve conservation science, natural resource

- management, and environmental protection. *Biological Conservation* 208:15–28. DOI: 10.1016/j.biocon.2016.05.015.
- McNeely JA. 1992. The sinking ark: pollution and the worldwide loss of biodiversity. *Biodiversity & Conservation* 1:2–18. DOI: 10.1007/BF00700247.
- Mela CF, Kopalle PK. 2002. The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations. *Applied Economics* 34:667–677. DOI: 10.1080/00036840110058482.
- Mesgaran MB, Cousens RD, Webber BL. 2014. Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Diversity and Distributions* 20:1147–1159. DOI: 10.1111/ddi.12209.
- Midgley M. 1994. Bridge-building at alst. In: *Animals and Human Society: Changing Perspectives*. London: Routledge,. DOI: 10.4324/9780203421444.
- Miller JR. 2005. Biodiversity conservation and the extinction of experience. *Trends in Ecology & Evolution* 20:430–434. DOI: 10.1016/j.tree.2005.05.013.
- Miller DL, Burt ML, Rexstad EA, Thomas L. 2013. Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution* 4:1001–1010. DOI: 10.1111/2041-210X.12105.
- Miller DAW, Weir LA, McClintock BT, Grant EHC, Bailey LL, Simons TR. 2012. Experimental investigation of false positive errors in auditory species occurrence surveys. *Ecological Applications* 22:1665–1674. DOI: 10.1890/11-2129.1.
- Minamoto T. 2022. Environmental DNA analysis for macro-organisms: species distribution and more. *DNA Research* 29:dsac018. DOI: 10.1093/dnares/dsac018.
- Mizel JD, Schmidt JH, Lindberg MS. 2018. Accommodating temporary emigration in spatial distance sampling models. *Journal of Applied Ecology* 55:1456–1464. DOI: 10.1111/1365-2664.13053.
- Monnet A-C, Jiguet F, Meynard CN, Mouillot D, Mouquet N, Thuiller W, Devictor V. 2014. Asynchrony of taxonomic, functional and phylogenetic diversity in birds. *Global Ecology and Biogeography* 23:780–788. DOI: 10.1111/geb.12179.
- Moolna A, Knott C, Wallis D, Crawshaw D, Brierley-Moore J, Simons J, Selby A. 2018. Engaging People with Carbon and Climate Change Using Landscape Scale Conservation and Biodiversity Monitoring. In: Leal Filho W, Manolas E, Azul AM, Azeiteiro UM, McGhie H eds. *Handbook of Climate Change Communication: Vol. 1: Theory of Climate Change Communication*. Climate Change Management. Cham: Springer International Publishing, 293–308. DOI: 10.1007/978-3-319-69838-0\_16.
- Morelli F, Brlík V, Benedetti Y, Bussière R, Moudrá L, Reif J, Svitok M. 2022. Detection Rate of Bird Species and What It Depends on: Tips for Field Surveys. *Frontiers in Ecology and Evolution* 9. DOI: 10.3389/fevo.2021.671492.
- Morin DJ, Yackulic CB, Diffendorfer JE, Lesmeister DB, Nielsen CK, Reid J, Schaub EM. 2020. Is your ad hoc model selection strategy affecting your multimodel inference? *Ecosphere* 11:e02997. DOI: 10.1002/ecs2.2997.
- Morris RF. 1955. The development of sampling techniques for forest insect defoliators, with particular reference to the spruce budworm. *Canadian Journal of Zoology* 33:225–294. DOI: 10.1139/z55-015.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Morris RF. 1960. Sampling Insect Populations. *Annual Review of Entomology* 5:243–264. DOI: 10.1146/annurev.en.05.010160.001331.
- Moss S. 2004a. 1. Observing: The First Birdwatchers. In: *A bird in the bush: a social history of birdwatching*. London: Aurum Press,.
- Moss S. 2004b. 9. Counting: Between the Wars, 1918-1939. In: *A bird in the bush: a social history of birdwatching*. London: Aurum Press,.
- Mouchet MA, Villéger S, Mason NWH, Mouillot D. 2010. Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. *Functional Ecology* 24:867–876. DOI: 10.1111/j.1365-2435.2010.01695.x.
- Mouillot D, Graham NAJ, Villéger S, Mason NWH, Bellwood DR. 2013. A functional approach reveals community responses to disturbances. *Trends in Ecology & Evolution* 28:167–177. DOI: 10.1016/j.tree.2012.10.004.
- Moussus J, Jiguet F, Clavel J, Julliard R. 2009. A method to estimate phenological variation using data from large-scale abundance monitoring programmes. *Bird Study* 56:198–212. DOI: 10.1080/00063650902792064.
- Murray DL, Bastille-Rousseau G, Beaty LE, Homseth ML, Row JR, Thomson DH. 2020. 2. From Research Hypothesis to Model Selection: A Strategy for Robust Inference in Population Ecology. In: Murray DL, Sandercock BK eds. *Population Ecology in Practice*. Hoboken: Wiley-Blackwell,.
- Naguib M, Wiley RH. 2001. Estimating the distance to a source of sound: mechanisms and adaptations for long-range communication. *Animal Behaviour* 62:825–837. DOI: 10.1006/anbe.2001.1860.
- Newbold T, Hudson LN, Hill SLL, Contu S, Lysenko I, Senior RA, Börger L, Bennett DJ, Choimes A, Collen B, Day J, De Palma A, Díaz S, Echeverria-Londoño S, Edgar MJ, Feldman A, Garon M, Harrison MLK, Alhousseini T, Ingram DJ, Itescu Y, Kattge J, Kemp V, Kirkpatrick L, Kleyer M, Correia DLP, Martin CD, Meiri S, Novosolov M, Pan Y, Phillips HRP, Purves DW, Robinson A, Simpson J, Tuck SL, Weiher E, White HJ, Ewers RM, Mace GM, Scharlemann JPW, Purvis A. 2015. Global effects of land use on local terrestrial biodiversity. *Nature* 520:45–50. DOI: 10.1038/nature14324.
- Newton I. 2023. 2. Methodology for migration studies. In: *The Migration Ecology of Birds*. San Diego: Academic Press,.
- Nichols JD, Hines JE, Sauer JR, Fallon FW, Fallon JE, Heglund PJ. 2000. A Double-Observer Approach for Estimating Detection Probability and Abundance From Point Counts. *The Auk* 117:393–408. DOI: 10.1093/auk/117.2.393.
- Nichols JD, Noon BR, Stokes SL, Hines JE. 1981. Remarks on the use of mark-recapture methodology in estimating avian population size. In: Ralph CJ, Scott JM eds. *Estimating numbers of terrestrial birds: Proceedings of an international symposium held at Asilomar, California, October 26-31, 1980*. Studies in Avian Biology. Cooper Ornithological Society, 121–136.
- Nichols JD, Thomas L, Conn PB. 2009. Inferences About Landbird Abundance from Count Data: Recent Advances and Future Directions. In: Thomson DL, Cooch EG, Conroy MJ eds. *Modeling Demographic Processes In Marked Populations*. Environmental and Ecological Statistics. Boston, MA: Springer US, 201–235. DOI: 10.1007/978-0-387-78151-8\_9.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Nilsson C, Dokter AM, Schmid B, Scacco M, Verlinden L, Bäckman J, Haase G, Dell'Omo G, Chapman JW, Leijnse H, Liechti F. 2018. Field validation of radar systems for monitoring bird migration. *Journal of Applied Ecology* 55:2552–2564. DOI: 10.1111/1365-2664.13174.
- Numminen E, Jokinen M, Lindén A, Vanhatalo J. 2023. Species ecology can bias population estimates. *Biological Conservation* 283:110115. DOI: 10.1016/j.biocon.2023.110115.
- Pacifici K, Simons TR, Pollock KH. 2008. EFFECTS OF VEGETATION AND BACKGROUND NOISE ON THE DETECTION PROCESS IN AUDITORY AVIAN POINT-COUNT SURVEYS - Efectos de la Vegetación y del Ruido de Fondo en el Proceso de Detección de Aves Mediante Registros Auditivos en Puntos de Conteo. *The Auk* 125:600–607. DOI: 10.1525/auk.2008.07078.
- Parnesian C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics* 37:637–669. DOI: 10.1146/annurev.ecolsys.37.091305.110100.
- Pauly D. 1995. Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology & Evolution* 10:430. DOI: 10.1016/S0169-5347(00)89171-5.
- Pauly D. 2008. Global fisheries: a brief review.
- Pavanato Julião H. 2021. Development of integrated distance sampling models. Thesis. University of Otago.
- Pearson K. 1901. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2:559–572. DOI: 10.1080/14786440109462720.
- Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Heip C, Höft R, Hurtt G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M. 2013. Essential Biodiversity Variables. *Science* 339:277–278. DOI: 10.1126/science.1229931.
- Peterson AT, Papeş M, Soberón J. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling* 213:63–72. DOI: 10.1016/j.ecolmodel.2007.11.008.
- Pilgrim S, Cullen-Unsworth L, Smith D, Pretty J. 2008. Ecological Knowledge is Lost in Wealthier Communities and Countries. *Environmental science & technology* 42:1004–9. DOI: 10.1021/es070837v.
- Poggiato G, Gaüzere P, Martinez-Almoyna C, Deschamps G, Renaud J, Violle C, Münkemüller T, Thuiller W. 2023. Predicting combinations of community mean traits using joint modelling. *Global Ecology and Biogeography* 32:1409–1422. DOI: 10.1111/geb.13706.
- Pollock KH. 1982. A Capture-Recapture Design Robust to Unequal Probability of Capture. *The Journal of Wildlife Management* 46:752–757. DOI: 10.2307/3808568.
- Pollock KH, Nichols JD, Karanth KU. 2012. 8. Estimating demographic parameters. In: Boitani L, Powell RA eds. *Carnivore ecology and conservation: a handbook of techniques*. Techniques in ecology and conservation series. Oxford ; New York: Oxford University Press,.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Pollock LJ, Tingley R, Morris WK, Golding N, O'Hara RB, Parris KM, Vesk PA, McCarthy MA. 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution* 5:397–406. DOI: <https://doi.org/10.1111/2041-210X.12180>.
- Porter JH. 2000. Scientific Databases. In: Michener WK, Brunt JW eds. *Ecological data: design, management, and processing*. Malden, MA: Blackwell Science,.
- Princé K, Zuckerberg B. 2015. Climate change in our backyards: the reshuffling of North America's winter bird communities. *Global Change Biology* 21:572–585. DOI: <https://doi.org/10.1111/gcb.12740>.
- Quaintenne G, Gaudard C, Béchet A, Benmergui M, Boutteaux J-J, Cadiou B, Camberlein P, Chapalain F, Croset F, Culioli J-M, Dalloyau S, Debout G, Dubois P, Dulac P, Flitti A, Gallien F, Gendre N, Girard O, Havet S, Vincent-Martin N. 2020. Les oiseaux nicheurs rares et menacés en France en 2016 et 2017. Rare and endangered breeding bird survey in France in 2016-2017. *Ornithos* 27-2:73–111.
- Rabinowitz D. 1981. 17. Seven forms of rarity. In: Syngé H, Linnean Society of London, King's College eds. *The biological aspects of rare plant conservation: proc. of an international Conference held at King's College, Cambridge, England, 14 - 19 July 1980*. A Wiley-interscience publication. Chichester: Wiley,.
- Ramachandran V, Robin VV, Tamma K, Ramakrishnan U. 2017. Climatic and geographic barriers drive distributional patterns of bird phenotypes within peninsular India. *Journal of Avian Biology* 48:620–630. DOI: 10.1111/jav.01278.
- Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A. 2006. Are niche-based species distribution models transferable in space? *Journal of Biogeography* 33:1689–1703. DOI: 10.1111/j.1365-2699.2006.01466.x.
- Rehme SE, Powell LA, Allen CR. 2011. Multimodel inference and adaptive management. *Journal of Environmental Management* 92:1360–1364. DOI: 10.1016/j.jenvman.2010.10.012.
- Réseau "Oiseaux de passage." 2017. Suivi des populations nicheuses (1996-2017) et hivernantes (2000-2017).
- Ricotta C, Bacaro G. 2010. On plot-to-plot dissimilarity measures based on species functional traits. *Community Ecology* 11:113–119. DOI: 10.1556/ComEc.11.2010.1.16.
- Roché J-E, Muller Y, Sibley J-P. 2013. Une méthode simple pour estimer les populations d'oiseaux communs nicheurs en France. *Alauda* 81:241–268.
- Roney HB. 1879. Efforts to Check the Slaughter. In: *The Passenger Pigeon*.
- Rosenberg AA, McLeod KL. 2005. Implementing ecosystem-based approaches to management for the conservation of ecosystem services. *Marine Ecology Progress Series* 300:270–274.
- Rosser AM, Mainka SA. 2002. Overexploitation and Species Extinctions. *Conservation Biology* 16:584–586.
- Royle JA, Dawson DK, Bates S. 2004. Modeling Abundance Effects in Distance Sampling. *Ecology* 85:1591–1597. DOI: 10.1890/03-3127.
- Royle JA, Dorazio RM. 2008. 1. Conceptual and Philosophical considerations in Ecology and Statistics. In: *Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities*.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Rue H, Martino S, Chopin N. 2009. Approximate Bayesian Inference for Latent Gaussian models by using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 71:319–392. DOI: 10.1111/j.1467-9868.2008.00700.x.
- Sánchez-Cordero V, Cirelli V, Mungui M, Sarkar S. 2005. Place prioritization for biodiversity content using species ecological niche modeling. *Biodiversity Informatics* 2. DOI: 10.17161/bi.v2i0.9.
- Santangeli A, Mammola S, Lehikoinen A, Rajasärkkä A, Lindén A, Saastamoinen M. 2022. The effects of protected areas on the ecological niches of birds and mammals. *Scientific Reports* 12:11601. DOI: 10.1038/s41598-022-15949-2.
- Sauer JR, Pardieck KL, Ziolkowski DJ Jr, Smith AC, Hudson M-AR, Rodriguez V, Berlanga H, Niven DK, Link WA. 2017. The first 50 years of the North American Breeding Bird Survey. *The Condor* 119:576–593. DOI: 10.1650/CONDOR-17-83.1.
- Schmidt BR, Cruickshank SS, Bühler C, Bergamini A. 2023. Observers are a key source of detection heterogeneity and biased occupancy estimates in species monitoring. *Biological Conservation* 283:110102. DOI: 10.1016/j.biocon.2023.110102.
- Schmidt JH, McIntyre CL, MacCluskie MC. 2013. Accounting for incomplete detection: What are we estimating and how might it affect long-term passerine monitoring programs? *Biological Conservation* 160:130–139. DOI: 10.1016/j.biocon.2013.01.007.
- Schopenhauer A. 1966. *The World as Will and Representation*. Courier Corporation.
- Schorger AW. 1955. 9. Decrease and extinction. In: *The passenger pigeon, its natural history and extinction*.
- Schwartz MW. 2012. Using niche models with climate projections to inform conservation management decisions. *Biological Conservation* 155:149–156. DOI: 10.1016/j.biocon.2012.06.011.
- Schwarz G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics* 6:461–464. DOI: 10.1214/aos/1176344136.
- Scott P, Burton JA, Fitter R. 1987. Red Data books: the historical background. In: Fitter M, Fitter RSR eds. *The road to extinction: problems of categorizing the status of taxa threatened with extinction: proceedings of a symposium held by the Species Survival Commission, Madrid, 7 and 9 November 1984*. IUCN,.
- Shannon CE. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27:379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.
- Simberloff D. 1998. Flagships, umbrellas, and keystones: Is single-species management passé in the landscape era? *Biological Conservation* 83:247–257. DOI: 10.1016/S0006-3207(97)00081-5.
- Slack NG, Wilson EO. 2011. 18. Concluding remarks: Hutchinson's Legacy in Ecology. In: G. Evelyn Hutchinson and the Invention of Modern Ecology. New Haven (Conn.): Yale University Press,.
- Snyder R, Clark M, Salas L, Schackwitz W, Leland D, Stephens T, Erickson T, Tuffli T, Tuffli M, Clas K. 2022. The Soundscapes to Landscapes Project: Development of a Bioacoustics-Based Monitoring Workflow with Multiple Citizen Scientist Contributions. *Citizen Science: Theory and Practice* 7. DOI: 10.5334/cstp.391.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Soga M, Gaston KJ. 2018. Shifting baseline syndrome: causes, consequences, and implications. *Frontiers in Ecology and the Environment* 16:222–230. DOI: 10.1002/fee.1794.
- Sólymos P, Matsuoka SM, Stralberg D, Barker NKS, Bayne EM. 2018. Phylogeny and species traits predict bird detectability. *Ecography* 41:1595–1603. DOI: 10.1111/ecog.03415.
- Soroye P, Newbold T, Kerr J. 2020. Climate change contributes to widespread declines among bumble bees across continents. *Science* 367:685–688. DOI: 10.1126/science.aax8591.
- Southwood T, Henderson P. 2000a. 1. Introduction to the Study of Animal Populations. In: *Ecological Methods* 3rd edition.
- Southwood T, Henderson P. 2000b. 9. Wildlife Population Estimates by Census and Distance-measuring Techniques. In: *Ecological Methods* 3rd edition.
- Stanton JC. 2014. Present-day risk assessment would have predicted the extinction of the passenger pigeon (*Ectopistes migratorius*). *Biological Conservation* 180:11–20. DOI: 10.1016/j.biocon.2014.09.023.
- Steffen W, Grinevald J, Crutzen P, McNeill J. 2011. The Anthropocene: conceptual and historical perspectives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 369:842–867. DOI: 10.1098/rsta.2010.0327.
- Stephens PA, Buskirk SW, del Rio CM. 2007. Inference in ecology and evolution. *Trends in Ecology & Evolution* 22:192–197. DOI: 10.1016/j.tree.2006.12.003.
- Strickfaden KM, Fagre DA, Golding JD, Harrington AH, Reintsma KM, Tack JD, Dreitz VJ. 2020. Dependent double-observer method reduces false-positive errors in auditory avian survey data. *Ecological Applications* 30:e02026. DOI: 10.1002/eap.2026.
- Sueur J, Farina A, Gasc A, Pieretti N, Pavoine S. 2014. Acoustic Indices for Biodiversity Assessment and Landscape Investigation. *Acta Acustica united with Acustica* 100:772–781. DOI: 10.3813/AAA.918757.
- Sullivan BL, Wood CL, Iliff MJ, Bonney RE, Fink D, Kelling S. 2009. eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation* 142:2282–2292. DOI: 10.1016/j.biocon.2009.05.006.
- Szabo JK, Fuller RA, Possingham HP. 2012. A comparison of estimates of relative abundance from a weakly structured mass-participation bird atlas survey and a robustly designed monitoring scheme. *Ibis* 154:468–479. DOI: 10.1111/j.1474-919X.2012.01229.x.
- Tannenbaum J, Bennett BT. 2015. Russell and Burch's 3Rs Then and Now: The Need for Clarity in Definition and Purpose. *Journal of the American Association for Laboratory Animal Science : JAALAS* 54:120–132.
- Thierion V, Vincent A, Valero S. 2022. Theia OSO Land Cover Map 2020. DOI: 10.5281/zenodo.6538861.
- Thompson WL. 2002. Towards Reliable Bird Surveys: Accounting for Individuals Present but not Detected. *The Auk* 119:18–25. DOI: 10.1642/0004-8038(2002)119[0018:TRBSAF]2.0.CO;2.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- Thompson WL, White GC, Gowan C. 1998a. 3. Enumeration methods. In: *Monitoring Vertebrate Populations*. San Diego: Academic Press,.
- Thompson WL, White GC, Gowan C. 1998b. 1. Basic Concepts. In: *Monitoring Vertebrate Populations*. San Diego: Academic Press,.
- Tilman D, Reich PB, Knops J, Wedin D, Mielke T, Lehman C. 2001. Diversity and Productivity in a Long-Term Grassland Experiment. *Science* 294:843–845. DOI: 10.1126/science.1060391.
- Travis JMJ. 2003. Climate change and habitat destruction: a deadly anthropogenic cocktail. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270:467–473. DOI: 10.1098/rspb.2002.2246.
- Tredennick AT, Hooker G, Ellner SP, Adler PB. 2021. A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology* 102:e03336. DOI: 10.1002/ecy.3336.
- Tylianakis JM, Didham RK, Bascompte J, Wardle DA. 2008. Global change and species interactions in terrestrial ecosystems. *Ecology Letters* 11:1351–1363. DOI: 10.1111/j.1461-0248.2008.01250.x.
- Tyre AJ, Tenhumberg B, Field SA, Niejalke D, Parris K, Possingham HP. 2003. Improving Precision and Reducing Bias in Biological Surveys: Estimating False-Negative Error Rates. *Ecological Applications* 13:1790–1801. DOI: 10.1890/02-5078.
- UICN France, MNHN, LPO, SEOF, ONCFS. 2016. *La Liste rouge des espèces menacées en France - Chapitre Oiseaux de France métropolitaine*. Paris, France.
- United Nations. 1993. Volume I : Resolutions Adopted by the Conference. In: *Report of the United Nations Conference on Environment and Development*. Rio de Janeiro,.
- Vallé C, Poggiato G, Thuiller W, Jiguet F, Princé K, Le Viol I. 2024. Species associations in joint species distribution models: from missing variables to conditional predictions. *Journal of Biogeography* 51:311–324. DOI: 10.1111/jbi.14752.
- Van Niekerk J, Krainski E, Rustand D, Rue H. 2023. A new avenue for Bayesian inference with INLA. *Computational Statistics & Data Analysis* 181:107692. DOI: 10.1016/j.csda.2023.107692.
- Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P-C. 2021. Rank-Normalization, Folding, and Localization: An Improved  $\hat{R}$  for Assessing Convergence of MCMC (with Discussion). *Bayesian Analysis* 16:667–718. DOI: 10.1214/20-BA1221.
- Verner J. 1985. Assessment of Counting Techniques. In: Johnston RF ed. *Current Ornithology: Volume 2*. Boston, MA: Springer US, 247–302. DOI: 10.1007/978-1-4613-2385-3\_8.
- Violle C, Navas M-L, Vile D, Kazakou E, Fortunel C, Hummel I, Garnier E. 2007. Let the concept of trait be functional! *Oikos* 116:882–892. DOI: 10.1111/j.0030-1299.2007.15559.x.
- Visser ME, Holleman LJM, Gienapp P. 2006. Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia* 147:164–172. DOI: 10.1007/s00442-005-0299-6.
- Warton DI, Blanchet FG, O'Hara RB, Ovaskainen O, Taskinen S, Walker SC, Hui FKC. 2015. So Many Variables: Joint Modeling in Community Ecology. *Trends in Ecology & Evolution* 30:766–779. DOI: 10.1016/j.tree.2015.09.007.

*Références de l'introduction générale Transmission de l'incertitude scientifique face à la prise d'action*

- White G. 1789. The natural history and antiquities of Selborne, in the county of Southampton : with engravings, and an appendix. London: Printed by T. Bensley, for B. White and Son, at Horace's Head, Fleet Street. DOI: 10.5962/bhl.title.104749.
- Whittaker RH. 1972. Evolution and Measurement of Species Diversity. *Taxon* 21:213–251. DOI: 10.2307/1218190.
- Whittaker RJ, Araújo MB, Jepson P, Ladle RJ, Watson JEM, Willis KJ. 2005. Conservation Biogeography: assessment and prospect. *Diversity and Distributions* 11:3–23. DOI: 10.1111/j.1366-9516.2005.00143.x.
- Wilhere GF. 2002. Adaptive Management in Habitat Conservation Plans. *Conservation Biology* 16:20–29. DOI: 10.1046/j.1523-1739.2002.00350.x.
- Wilson A. 1829. Passenger Pigeon. In: *American ornithology; or, The natural history of the birds of the United States...* New York: Harrison Hall, 1–424. DOI: 10.5962/bhl.title.25877.
- Wilson KA, Law EA. 2016. Ethics of Conservation Triage. *Frontiers in Ecology and Evolution* 4. DOI: 10.3389/fevo.2016.00112.
- Wood SN. 2017. *Generalized Additive Models : An Introduction with R, Second Edition.* Chapman and Hall/CRC. DOI: 10.1201/9781315370279.
- Wu J. 2013. Hierarchy Theory: An Overview. In: Rozzi R, Pickett STA, Palmer C, Armesto JJ, Callicott JB eds. *Linking Ecology and Ethics for a Changing World: Values, Philosophy, and Action.* Dordrecht: Springer Netherlands, 281–301. DOI: 10.1007/978-94-007-7470-4\_24.
- WWF. 2020. *Living Planet Report 2020. Bending the Curve of Biodiversity Loss: A Deep Dive into the Living Planet Index.* Gland: World Wildlife Fund.
- Yeatman L, Ridet Y, d'ornithologie S française, France D de la protection de la nature. 1976. *Atlas des oiseaux nicheurs de France de 1970 à 1975.* Paris, France: Société ornithologique de France.
- Yeatman-Berthelot D, Jarry G. 1995. *Nouvel atlas des oiseaux nicheurs de France 1985-1989.* Paris: Societe Ornithologique De Fran.
- Yoccoz NG, Nichols JD, Boulinier T. 2001. Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution* 16:446–453. DOI: 10.1016/S0169-5347(01)02205-4.
- Zavaleta E, Heller N. 2012. Responses of Communities and Ecosystems to Global Changes. In: Levin SA, Carpenter SR, Godfray HCJ, Kinzig AP, Loreau M, Losos JB, Walker B, Wilcove DS eds. *The Princeton Guide to Ecology.* Princeton, NJ: Princeton University Press,.
- Zemanova MA. 2020. Towards more compassionate wildlife research through the 3Rs principles: moving from invasive to non-invasive methods. *Wildlife Biology* 2020:wlb.00607. DOI: 10.2981/wlb.00607.
- Zipkin EF, Zylstra ER, Wright AD, Saunders SP, Finley AO, Dietze MC, Itter MS, Tingley MW. 2021. Addressing data integration challenges to link ecological processes across scales. *Frontiers in Ecology and the Environment* 19:30–38. DOI: 10.1002/fee.2290.



## VIII Références de la discussion générale

---

- Abadi F, Gimenez O, Arlettaz R, Schaub M. 2010. An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. *Ecology* 91:7–14. DOI: 10.1890/08-2235.1.
- Almond REA, Grooten M, Juffe B, Peterson T. 2022. *Living Planet Report 2022 – Building a naturepositive society*. WWF.
- Beck S, Mahdad M, Beukel K, Poetz M. 2019. The Value of Scientific Knowledge Dissemination for Scientists—A Value Capture Perspective. *Publications* 7:54. DOI: 10.3390/publications7030054.
- Besbeas P, Freeman SN, Morgan BJT, Catchpole EA. 2002. Integrating Mark-Recapture-Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters. *Biometrics* 58:540–547.
- Bodoni S. 2017. Bayer, Syngenta Clash With EU Over Bees Amid M&A Charm Offensive. *Bloomberg.com*.
- Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips T, Rosenberg KV, Shirk J. 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience* 59:977–984. DOI: 10.1525/bio.2009.59.11.9.
- Brownell SE, Price JV, Steinman L. 2013. Science Communication to the General Public: Why We Need to Teach Undergraduate and Graduate Students this Skill as Part of Their Formal Scientific Training. *Journal of Undergraduate Neuroscience Education* 12:E6–E10.
- Buxton RT, Avery-Gomm S, Lin H-Y, Smith PA, Cooke SJ, Bennett JR. 2020. Half of resources in threatened species conservation plans are allocated to research and monitoring. *Nature Communications* 11:4668. DOI: 10.1038/s41467-020-18486-6.
- Buxton RT, Nyboer EA, Pigeon KE, Raby GD, Rytwinski T, Gallagher AJ, Schuster R, Lin H-Y, Fahrig L, Bennett JR, Cooke SJ, Roche DG. 2021. Avoiding wasted research resources in conservation science. *Conservation Science and Practice* 3:e329. DOI: 10.1111/csp2.329.
- Callaghan CT, Poore AGB, Major RE, Rowley JLL, Cornwell WK. 2019. Optimizing future biodiversity sampling by citizen scientists. *Proceedings of the Royal Society B: Biological Sciences* 286:20191487. DOI: 10.1098/rspb.2019.1487.
- Castagneyrol B, Bedessem B, Julliard R. 2023. Is ecology different when studied with citizen scientists? A bibliometric analysis. *Ecology and Evolution* 13:e10488. DOI: 10.1002/ece3.10488.
- Chandler M, See L, Copas K, Bonde AMZ, López BC, Danielsen F, Legind JK, Masinde S, Miller-Rushing AJ, Newman G, Rosemartin A, Turak E. 2017. Contribution of citizen science towards international biodiversity monitoring. *Biological Conservation* 213:280–294. DOI: 10.1016/j.biocon.2016.09.004.

- Collen B, Loh J, Whitmee S, McRAE L, Amin R, Baillie JEM. 2009. Monitoring Change in Vertebrate Abundance: the Living Planet Index. *Conservation Biology* 23:317–327. DOI: 10.1111/j.1523-1739.2008.01117.x.
- Cooke S, Gallagher A, Sopinka N, Nguyen V, Skubel R, Hammerschlag N, Boon S, Young N, Danylchuk A. 2017. Considerations for effective science communication. *FACETS* 2. DOI: 10.1139/facets-2016-0055.
- Cooney R. 2005. 2. The meaning of the precautionary principle. In: *Biodiversity and the Precautionary Principle: Risk and Uncertainty in Conservation and Sustainable Use*.
- Courchamp F, Angulo E, Rivalan P, Hall RJ, Signoret L, Bull L, Meinard Y. 2006. Rarity Value and Species Extinction: The Anthropogenic Allee Effect. *PLOS Biology* 4:e415. DOI: 10.1371/journal.pbio.0040415.
- Daskalova GN, Myers-Smith IH, Godlee JL. 2020. Rare and common vertebrates span a wide spectrum of population trends. *Nature Communications* 11:4394. DOI: 10.1038/s41467-020-17779-0.
- Desrosières A. 2003. Les qualités des quantités. In: *Courrier des statistiques n° 105-106*.
- Deutz A, Heal GM, Niu R, Swanson E, Townsend T, Zhu L, Delmar A, Meghji A, Sethi SA, Tobin de la Puente J. 2020. *Financing Nature: Closing the global biodiversity financing gap*. The Paulson Institute, The Nature Conservancy, and the Cornell Atkinson Center for Sustainability.
- Devictor V, Whittaker RJ, Beltrame C. 2010. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions* 16:354–362. DOI: 10.1111/j.1472-4642.2009.00615.x.
- Dove S, Böhm M, Freeman R, McRae L, Murrell DJ. 2023. Quantifying reliability and data deficiency in global vertebrate population trends using the Living Planet Index. *Global Change Biology* 29:4966–4982. DOI: 10.1111/gcb.16841.
- Efron B. 2012. A 250-year argument: Belief, behavior, and the bootstrap. *Bulletin of the American Mathematical Society* 50:129–146. DOI: 10.1090/S0273-0979-2012-01374-5.
- Emlen JT. 1971. Population Densities of Birds Derived from Transect Counts. *The Auk* 88:323–342. DOI: 10.2307/4083883.
- Fink D, Auer T, Johnston A, Ruiz-Gutierrez V, Hochachka WM, Kelling S. 2020. Modeling avian full annual cycle distribution and population trends with citizen science data. *Ecological Applications* 30:e02056. DOI: 10.1002/eap.2056.
- Fontaine B, Moussy C, Carricaburu JC, Dupuis J, Schmaltz L, Lorrillière R, Lois G, Gaudard C. 2020. *Suivi des oiseaux communs en France 1989-2019 : 30 ans de suivis participatifs*.
- Fontaine A, Simard A, Brunet N, Elliott KH. 2022. Scientific contributions of citizen science applied to rare or threatened animals. *Conservation Biology* 36:e13976. DOI: 10.1111/cobi.13976.
- Gould E, Fraser HS, Parker TH, Nakagawa S, Griffith SC, Vesk PA, Fidler F, Hamilton DG, Abbey-Lee RN, Abbott JK, Aguirre LA, Alcaraz C, Aloni I, Altschul D, Arekar K, Atkins JW, Atkinson J, Baker C, Barrett M, Bell K, Bello SK, Beltrán I, Berauer BJ, Bertram MG, Billman PD, Blake CK, Blake S, Bliard L, Bonisoli-Alquati A, Bonnet T, Bordes CNM, Bose APH, Botterill-James T, Boyd MA, Boyle SA, Bradfer-Lawrence T, Bradham J, Brand JA, Brengdahl MI, Bulla M, Bussière L, Camerlenghi E, Campbell

SE, Campos LLF, Caravaggi A, Cardoso P, Carroll CJW, Catanach TA, Chen X, Chik HYJ, Choy ES, Christie AP, Chuang A, Chunco AJ, Clark BL, Contina A, Covernton GA, Cox MP, Cressman KA, Crotti M, Crouch CD, D'Amelio PB, Sousa AA de, Döbert TF, Dobler R, Dobson AJ, Doherty TS, Drobniak SM, Duffy AG, Duncan AB, Dunn RP, Dunning J, Dutta T, Eberhart-Hertel L, Elmore JA, Elsherif MM, English HM, Ensminger DC, Ernst UR, Ferguson SM, Fernández-Juricic E, Ferreira-Arruda T, Fieberg J, Finch EA, Fiorenza EA, Fisher DN, Fontaine A, Forstmeier W, Fourcade Y, Frank GS, Freund CA, Fuentes-Lillo E, Gandy SL, Gannon DG, García-Cervigón AI, Garretson AC, Ge X, Geary WL, Géron C, Gilles M, Girndt A, Glikzman D, Goldspiel HB, Gomes DGE, Good MK, Goslee SC, Gosnell JS, Grames EM, Gratton P, Grebe NM, Greenler SM, Griffioen M, Griffith DM, Griffith FJ, Grossman JJ, Güncan A, Haesen S, Hagan JG, Hager HA, Harris JP, Harrison ND, Hasnain SS, Havird JC, Heaton AJ, Herrera-Chaustre ML, Howard TJ, Hsu B-Y, Iannarilli F, Iranzo EC, Iverson ENK, Jimoh SO, Johnson DH, Johnsson M, Jorna J, Jucker T, Jung M, Kačergytė I, Kaltz O, Ke A, Kelly CD, Keogan K, Keppeler FW, Killion AK, Kim D, Kochan DP, Korsten P, Kothari S, Kuppler J, Kusch JM, Lagisz M, Lalla KM, Larkin DJ, Larson CL, Lauck KS, Lauterbur ME, Law A, Léandri-Breton D-J, Lembrechts JJ, L'Herpinierie K, Lievens EJP, Lima DO de, Lindsay S, Luquet M, MacLeod R, Macphie KH, Magellan K, Mair MM, Malm LE, Mammola S, Mandeville CP, Manhart M, Manrique-Garzon LM, Mäntylä E, Marchand P, Marshall BM, Martin CA, Martin DA, Martin JM, Martinig AR, McCallum ES, McCauley M, McNew SM, Meiners SJ, Merklung T, Michelangeli M, Moiron M, Moreira B, Mortensen J, Mos B, Muraina TO, Murphy PW, Nelli L, Niemelä P, Nightingale J, Nilsonne G, Nolazco S, Nooten SS, Novotny JL, Olin AB, Organ CL, Ostevik KL, Palacio FX, Paquet M, Parker DJ, Pascall DJ, Pasquarella VJ, Paterson JH, Payo-Payo A, Pedersen KM, Perez G, Perry KI, Pottier P, Proulx MJ, Proulx R, Pruett JL, Ramananjato V, Randimbiarison FT, Razafindratsima OH, Rennison DJ, Riva F, Riyahi S, Roast MJ, Rocha FP, Roche DG, Román-Palacios C, Rosenberg MS, Ross J, Rowland FE, Rugemalila D, Russell AL, Ruuskanen S, Saccone P, Sadeh A, Salazar SM, Sales K, Salmón P, Sánchez-Tójar A, Santos LP, Santostefano F, Schilling HT, Schmidt M, Schmoll T, Schneider AC, Schrock AE, Schroeder J, Schtickzelle N, Schultz NL, Scott DA, Scroggie MP, Shapiro JT, Sharma N, Shearer CL, Simón D, Sitvarin MI, Skupien FL, Slinn HL, Smith GP, Smith JA, Sollmann R, Whitney KS, Still SM, Stuber EF, Sutton GF, Swallow B, Taff CC, Takola E, Tanentzap AJ, Tarjuelo R, Telford RJ, Thawley CJ, Thierry H, Thomson J, Tidau S, Tompkins EM, Tortorelli CM, Trlica A, Turnell BR, Urban L, Vondel SV de, Wal JEM van der, Eeckhoven JV, Oordt F van, Vanderwel KM, Vanderwel MC, Vanderwolf KJ, Vélez J, Vergara-Florez DC, Verrelli BC, Vieira MV, Villamil N, Vitali V, Vollering J, Walker J, Walker XJ, Walter JA, Waryszak P, Weaver RJ, Wedegärtner REM, Weller DL, Whelan S, White RL, Wolfson DW, Wood A, Yanco SW, Yen JDL, Youngflesh C, Zilio G, Zimmer C, Zimmerman GM, Zitomer RA. 2023. Same data, different analysts: variation in effect sizes due to analytical decisions in ecology and evolutionary biology.

Herweijer C, Evison W, Mariam S, Khatri A, Albani M, Semov A, Long E. 2020. *Nature Risk Rising: Why the Crisis Engulfing Nature Matters for Business and the Economy*.

- IPBES Technical Support Unit on Knowledge and. 2021. IPBES regions and sub-regions. DOI: 10.5281/zenodo.5719431.
- Isaac NJB, van Strien AJ, August TA, de Zeeuw MP, Roy DB. 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* 5:1052–1060. DOI: 10.1111/2041-210X.12254.
- Jiguet F, Arlettaz R, Bauer H-G, Belik V, Copete JL, Couzi L, Czajkowski MA, Dale S, Dombrovski V, Elts J, Ferrand Y, Hargues R, Kirwan GM, Minkevicius S, Piha M, Selstam G, Skierczyński M, Sibley J-P, Sokolov A. 2016. An update of the European breeding population sizes and trends of the Ortolan Bunting (*Emberiza hortulana*). *Ornis Fennica* 93:186–196.
- Jiguet F, Robert A, Lorrillière R, Hobson KA, Kardynal KJ, Arlettaz R, Bairlein F, Belik V, Bernardy P, Copete JL, Czajkowski MA, Dale S, Dombrovski V, Ducros D, Efrat R, Elts J, Ferrand Y, Marja R, Minkevicius S, Olsson P, Pérez M, Piha M, Raković M, Schmaljohann H, Seimola T, Selstam G, Sibley J-P, Skierczyński M, Sokolov A, Sondell J, Moussy C. 2019. Unravelling migration connectivity reveals unsustainable hunting of the declining ortolan bunting. *Science Advances* 5:eau2642. DOI: 10.1126/sciadv.aau2642.
- Johnston A, Hochachka WM, Strimas-Mackey ME, Gutierrez VR, Robinson OJ, Miller ET, Auer T, Kelling ST, Fink D. 2019. Best practices for making reliable inferences from citizen science data: case study using eBird to estimate species distributions. *bioRxiv*:574392. DOI: 10.1101/574392.
- Johnston A, Matechou E, Dennis EB. 2022. Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution* 14:103–116. DOI: 10.1111/2041-210X.13834.
- Julliard R, Jiguet F. 2002. Un suivi intégré des populations d'oiseaux communs en France. *Alauda*:137–147.
- Kissel AM, Palen WJ, Ryan ME, Adams MJ. 2019. Compounding effects of climate change reduce population viability of a montane amphibian. *Ecological Applications* 29:e01832. DOI: 10.1002/eap.1832.
- Kosmala M, Wiggins A, Swanson A, Simmons B. 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment* 14:551–560. DOI: 10.1002/fee.1436.
- Ladle RJ, Alves-Martins F, Malhado ACM, Reyes-García V, Courchamp F, Minin ED, Roll U, Jarić I, Correia RA. 2023. Biocultural aspects of species extinctions. *Cambridge Prisms: Extinction* 1:e22. DOI: 10.1017/ext.2023.20.
- Ladle R, Correia R, Do Y, Joo G-J, Malhado A, Proulx R, Roberge J-M, Jepson P. 2016. Conservation culturomics. *Frontiers in Ecology and the Environment* 14:269–275. DOI: 10.1002/fee.1260.
- Ledger SEH, Loh J, Almond R, Böhm M, Clements CF, Currie J, Deinet S, Galewski T, Grooten M, Jenkins M, Marconi V, Painter B, Scott-Gatty K, Young L, Hoffmann M, Freeman R, McRae L. 2023. Past, present, and future of the Living Planet Index. *npj Biodiversity* 2:1–13. DOI: 10.1038/s44185-023-00017-3.
- Legendre S. 2020. 8. Projecting Populations. In: Murray DL, Sandercock BK eds. *Population Ecology in Practice*. Hoboken: Wiley-Blackwell.
- Levrel H. 2007. Quels indicateurs pour la gestion de la biodiversité ?

- LPO France. 2024. Le gouvernement suspend, mais toujours pour un an seuleme...
- Luhmann N. 1998. The Ecology of Ignorance. In: *Observations on modernity*. Writing science. Stanford, CA: Stanford University Press.,
- Mace PM, Sissenwine MP. 2002. Coping with Uncertainty: Evolution of the Relationship between Science and Management. *COPING WITH UNCERTAINTY*.
- Marconi V, McRae L, Müller H, Currie J, Whitmee S, Gadallah F (ZuZu), Freeman R. 2021. Population declines among Canadian vertebrates: But data of different quality show diverging trends. *Ecological Indicators* 130:108022. DOI: 10.1016/j.ecolind.2021.108022.
- Matutini F, Baudry J, Pain G, Sineau M, Python J. 2021. How citizen science could improve species distribution models and their independent assessment. *Ecology and Evolution* 11:3028–3039. DOI: 10.1002/ece3.7210.
- Maxim L, Van Der Sluijs JP. 2007. Uncertainty: Cause or effect of stakeholders' debates? *Science of The Total Environment* 376:1–17. DOI: 10.1016/j.scitotenv.2006.12.052.
- McDonald-Madden E, Baxter PWJ, Fuller RA, Martin TG, Game ET, Montambault J, Possingham HP. 2010. Monitoring does not always count. *Trends in Ecology & Evolution* 25:547–550. DOI: 10.1016/j.tree.2010.07.002.
- McKinley DC, Miller-Rushing AJ, Ballard HL, Bonney R, Brown H, Cook-Patton SC, Evans DM, French RA, Parrish JK, Phillips TB, Ryan SF, Shanley LA, Shirk JL, Stepenuck KF, Weltzin JF, Wiggins A, Boyle OD, Briggs RD, Chapin SF, Hewitt DA, Preuss PW, Soukup MA. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation* 208:15–28. DOI: 10.1016/j.biocon.2016.05.015.
- McRae L, Deinet S, Freeman R. 2017. The Diversity-Weighted Living Planet Index: Controlling for Taxonomic Bias in a Global Biodiversity Indicator. *PLOS ONE* 12:e0169156. DOI: 10.1371/journal.pone.0169156.
- Moreno J, Møller AP. 2011. Extreme climatic events in relation to global change and their impact on life histories. *Current Zoology* 57:375–389. DOI: 10.1093/czoolo/57.3.375.
- Mouterde P. 2020. Les populations de vertébrés ont chuté de 68 % en moins de cinquante ans.
- Norton JD. 1994. Science and certainty.
- Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, Buck S, Chambers CD, Chin G, Christensen G, Contestabile M, Dafoe A, Eich E, Freese J, Glennerster R, Goroff D, Green DP, Hesse B, Humphreys M, Ishiyama J, Karlan D, Kraut A, Lupia A, Mabry P, Madon T, Malhotra N, Mayo-Wilson E, McNutt M, Miguel E, Paluck EL, Simonsohn U, Soderberg C, Spellman BA, Turitto J, VandenBos G, Vazire S, Wagenmakers EJ, Wilson R, Yarkoni T. 2015. Promoting an open research culture. *Science* 348:1422–1425. DOI: 10.1126/science.aab2374.
- Parker TH, Forstmeier W, Koricheva J, Fidler F, Hadfield JD, Chee YE, Kelly CD, Gurevitch J, Nakagawa S. 2016. Transparency in Ecology and Evolution: Real Problems, Real Solutions. *Trends in Ecology & Evolution* 31:711–719. DOI: 10.1016/j.tree.2016.07.002.
- Puurtinen M, Elo M, Kotiaho JS. 2022. The Living Planet Index does not measure abundance. *Nature* 601:E14–E15. DOI: 10.1038/s41586-021-03708-8.

- Quéroué M, Barbraud C, Barraquand F, Turek D, Delord K, Pacoureau N, Gimenez O. 2021. Multispecies integrated population model reveals bottom-up dynamics in a seabird predator–prey system. *Ecological Monographs* 91:e01459. DOI: 10.1002/ecm.1459.
- Quigley TP, Amdam GV, Harwood GH. 2019. Honey bees as bioindicators of changing global agricultural landscapes. *Current Opinion in Insect Science* 35:132–137. DOI: 10.1016/j.cois.2019.08.012.
- Reyes-García V, Fernández-Llamazares Á, Aumeeruddy-Thomas Y, Benyei P, Bussmann RW, Diamond SK, García-del-Amo D, Guadilla-Sáez S, Hanazaki N, Kosoy N, Lavidés M, Luz AC, McElwee P, Meretsky VJ, Newberry T, Molnár Z, Ruiz-Mallén I, Salpeteur M, Wyndham FS, Zorondo-Rodriguez F, Brondizio ES. 2022. Recognizing Indigenous peoples' and local communities' rights and agency in the post-2020 Biodiversity Agenda. *Ambio* 51:84–92. DOI: 10.1007/s13280-021-01561-7.
- Rigal S, Devictor V, Dakos V. 2020. A method for classifying and comparing non-linear trajectories of ecological variables. *Ecological Indicators* 112:106113. DOI: 10.1016/j.ecolind.2020.106113.
- Robinson RA, Julliard R, Saracco JF. 2009. Constant effort: Studying avian population processes using standardised ringing. *Ringing & Migration* 24:199–204. DOI: 10.1080/03078698.2009.9674392.
- Robinson OJ, Ruiz-Gutierrez V, Fink D, Meese RJ, Holyoak M, Cooch EG. 2018. Using citizen science data in integrated population models to inform conservation. *Biological Conservation* 227:361–368. DOI: 10.1016/j.biocon.2018.10.002.
- Schaub M. 2020. 9. Combining Counts of Unmarked Individuals and Demographic Data Using Integrated Population Models. In: Murray DL, Sandercock BK eds. *Population Ecology in Practice*. Hoboken: Wiley-Blackwell, 215–236.
- Schaub M, Abadi F. 2010. Integrated population models: a novel analysis framework for deeper insights into population dynamics. *J. Ornithol* 152:227–237. DOI: 10.1007/s10336-010-0632-7.
- Sun CC, Royle JA, Fuller AK. 2019. Incorporating citizen science data in spatially explicit integrated population models. *Ecology* 100:e02777. DOI: 10.1002/ecy.2777.
- Thompson WL. 2002. Towards Reliable Bird Surveys: Accounting for Individuals Present but not Detected. *The Auk* 119:18–25. DOI: 10.1642/0004-8038(2002)119[0018:TRBSAF]2.0.CO;2.
- Toszogyova A, Smyčka J, Storch D. 2024. Mathematical biases in the calculation of the Living Planet Index lead to overestimation of vertebrate population decline. *Nature Communications* 15:5295. DOI: 10.1038/s41467-024-49070-x.
- UICN France, MNHN, LPO, SEOF, ONCFS. 2016. *La Liste rouge des espèces menacées en France - Chapitre Oiseaux de France métropolitaine*. Paris, France.
- Van De Pol M, Vindenes Y, Sæther B-E, Engen S, Ens BJ, Oosterbeek K, Tinbergen JM. 2010. Effects of climate change and variability on population dynamics in a long-lived shorebird. *Ecology* 91:1192–1204. DOI: 10.1890/09-0410.1.
- Viana DS, Santoro S, Soriguer RC, Figuerola J. 2023. A synthesis of Eurasian Curlew (*Numenius arquata arquata*) demography and population viability to inform its management. *Ibis* 165:767–780. DOI: 10.1111/ibi.13184.

- van der Vliet R, Valluerca I, Dijk J, Wassen M. 2015. EU protection is inadequate for a declining flyway population of Black-tailed Godwit *Limosa limosa*: Mismatch between future core breeding areas and existing Special Protection Areas. *Bird Conservation International* 25:111–125. DOI: 10.1017/S0959270914000100.
- Whittaker RJ, Ladle RJ. 2011. 1. The roots of conservation biogeography. In: Ladle RJ, Whittaker RJ eds. *Conservation Biogeography*. John Wiley & Sons,.
- Wiens JA. 1989. Spatial Scaling in Ecology. *Functional Ecology* 3:385–397. DOI: 10.2307/2389612.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3:160018. DOI: 10.1038/sdata.2016.18.



## IX Annexe: Tableau des tailles de populations d'oiseaux communs nicheurs de France métropolitaine

---

Ce tableau correspond à une version allégée et traduite de l'appendice S6.1 de l'article 1 (*Supplementary information*). A noter, qu'initialement les estimations des effectifs nicheurs de l'atlas des oiseaux de France métropolitaine (AOFM ; [Issa & Muller, 2015](#)) sont données sous la forme de bornes minimales et maximales. Afin de comparer les effectifs estimés par les deux méthodes, j'ai utilisé la médiane des valeurs minimales et maximales des estimations de l'AOFM (cf. *Comparison of ArGeom and HDS estimated population sizes*).

Les estimations sont présentées en nombre de couples ou paires potentielles. Pour les estimations provenant de la campagne EPOC-ODF 2021-2023, en amont de la calibration des modèles, j'ai appliqué un filtre selon la possibilité d'identification des individus mâles (visuelle ou sonore) durant point d'écoute. Voir l'appendice S2.1 de l'article 1.

Les huit espèces en italiques, soulignées en rouge correspondent aux espèces ayant des variations d'estimations nationales importantes entre les estimations brutes (sans traitement des valeurs extrêmes) et les estimations après traitement (cf. *Trimming of HDS population size estimate: assessment of model extrapolation*)

Espèces	Statut de conservation	AOFM		EPOC-ODF 2021-2023			$\delta_{\text{mean}}$ (%)
		Estimation médiane		Estimation			
		2012	2023	Basse	Moyenne	Haute	
Fauvette a tete noire	LC	6.5M	7.1M	8.6M	9.5M	10.5M	-0.29
Rougegorge familier	LC	4.5M	7.2M	7.7M	8.5M	9.4M	-0.17
Pinson des arbres	LC	9.0M	8.3M	7.2M	7.9M	8.8M	0.04
Mesange bleue	LC	3.2M	3.5M	4.7M	5.4M	6.3M	-0.42
Mesange charbonniere	LC	5.5M	5.4M	4.9M	5.4M	6.1M	0.00
Merle noir	LC	6.5M	6.1M	4.0M	4.5M	5.1M	0.29
Chardonneret elegant	VU	1.5M	1.7M	3.0M	4.1M	5.5M	-0.82
Pigeon ramier	LC	2.5M	3.7M	3.2M	3.6M	3.9M	0.02
Pouillot veloce	LC	3.2M	3.2M	3.1M	3.5M	4.0M	-0.09
Troglodyte mignon	LC	4.0M	3.4M	3.0M	3.4M	3.9M	0.00
Hirondelle rustique	NT	1.3M	1.2M	1.4M	2.0M	2.9M	-0.51
Linotte melodieuse	VU	750.0K	1.0M	551.3K	1.7M	5.6M	-0.53
<i>Alouette des champs</i>	NT	1.6M	1.5M	1.5M	1.5M	2.1M	-0.03
Grimpereau des jardins	LC	1.2M	1.4M	1.1M	1.3M	1.6M	0.05
Roitelet a triple bandeau	LC	700.0K	809.2K	962.6K	1.3M	1.8M	-0.46
Accenteur mouchet	LC	1.5M	1.1M	895.6K	1.1M	1.5M	0.02
Grive musicienne	LC	1.5M	1.6M	966.4K	1.1M	1.2M	0.36
Bergeronnette grise	LC	550.0K	523.0K	713.4K	1.0M	1.5M	-0.66
Sittelle torchepot	LC	1.5M	1.0M	831.3K	1.0M	1.3M	0.00
<i>Roitelet huppe</i>	NT	650.0K	350.4K	1.3M	980.6K	2.7M	-0.95
<i>Bruant zizi</i>	LC	300.0K	344.1K	778.7K	968.5K	1.4M	-0.95
Rougequeue noir	LC	950.0K	855.0K	717.1K	966.4K	1.3M	-0.12
Pic epeiche	LC	550.0K	632.5K	752.6K	930.2K	1.2M	-0.38
Verdier Europe	VU	1.5M	862.5K	666.0K	881.2K	1.2M	-0.02
Rossignol philomele	LC	750.0K	714.8K	739.8K	874.9K	1.0M	-0.20
Grosbec cassenoiaux	LC	80.0K	66.0K	422.3K	747.4K	1.4M	-1.68
Tarier patre	NT	250.0K	307.8K	494.2K	730.1K	1.1M	-0.81
Serin cini	VU	375.0K	361.1K	532.4K	714.8K	979.5K	-0.66
Mesange nonnette	LC	800.0K	418.4K	440.1K	672.8K	1.1M	-0.47
Fauvette grisette	LC	1.0M	944.0K	475.9K	613.5K	795.1K	0.42
Epervier Europe	LC	47.3K	41.3K	30.2K	603.5K	7.3M	-1.74
Hypolais polyglotte	LC	675.0K	553.5K	379.2K	499.1K	662.2K	0.10
<i>Mesange huppee</i>	LC	450.0K	321.8K	312.2K	492.5K	840.5K	-0.42
<i>Bruant jaune</i>	VU	750.0K	379.5K	370.4K	453.6K	698.9K	-0.18
<i>Bouvreuil pivoine</i>	VU	150.0K	78.0K	269.1K	437.2K	1.1M	-1.39
Piegrieche ecorcheur	NT	150.0K	224.1K	244.1K	419.9K	752.0K	-0.61
Buse variable	LC	160.0K	152.3K	270.6K	367.8K	502.2K	-0.83

*Annexe: Tableau des tailles de populations d'oiseaux communs nicheurs de France métropolitaine*

Tourterelle des bois	VU	400.0K	296.4K	265.3K	344.1K	449.8K	-0.15	
Grive draine	LC	400.0K	386.4K	296.2K	341.4K	417.2K	0.12	
Pouillot de Bonelli	LC	225.0K	241.0K	148.8K	322.5K	757.7K	-0.29	
Loriot Europe	LC	150.0K	173.4K	244.9K	317.7K	409.3K	-0.59	
Pic vert	LC	225.0K	207.5K	263.2K	316.1K	380.8K	-0.42	
Faisan de Colchide	LC	205.0K	201.3K	196.1K	293.4K	434.3K	-0.37	
<i>Bergeronnette printanière</i>	LC	135.0K	116.5K	276.0K	249.9K	410.5K	-0.73	
Pipit des arbres	LC	375.0K	229.1K	105.4K	195.3K	355.6K	0.16	
Bruant proyer	LC	300.0K	363.0K	150.1K	194.4K	254.3K	0.60	
Alouette lulu	LC	140.0K	265.4K	153.1K	192.9K	245.2K	0.32	
Faucon crecerelle	NT	76.0K	81.5K	126.1K	192.5K	295.9K	-0.81	
Coucou gris	LC	225.0K	208.3K	129.8K	162.3K	199.9K	0.25	
Rougequeue a front blanc	LC	120.0K	137.5K	39.3K	129.8K	283.0K	0.06	
Pouillot fitis	NT	150.0K	83.7K	34.5K	113.1K	341.6K	-0.30	
Pic noir	LC	32.5K	38.6K	26.6K	44.0K	71.9K	-0.13	
<i>Pigeon colombin</i>	LC	45.0K	91.2K	20.3K	42.0K	86.6K	0.74	
Huppe fasciee	LC	85.0K	77.9K	25.3K	41.1K	64.5K	0.62	