

Mémoire de fin d'études

**Analyse de la performance et des biais liés à la reconnaissance automatisée
des chants par BirdNET à partir d'enregistrements acoustiques pour
l'étude des communautés d'oiseaux en milieu urbain**



Héloïse JEUX

Mars – Septembre 2024

Cursus Ingénieur AgroParisTech - Dominante Gestion des Milieux Naturels





Site 78 du relevé Evolville – 12 avril 2024, Strasbourg ©Héloïse Jeux

Mémoire de fin d'études

**Analyse de la performance et des biais liés à la reconnaissance automatisée
des chants par BirdNET à partir d'enregistrements acoustiques pour
l'étude des communautés d'oiseaux en milieu urbain**

Héloïse JEUX

Mars – Septembre 2024

Cursus Ingénieur AgroParisTech - Dominante Gestion des Milieux Naturels

Organisme du stage : Laboratoire Image Ville Environnement de l'Université de Strasbourg

Encadrement : Alejandro Sotillo (LIVE), François Chiron (LIVE)

FICHE SIGNALÉTIQUE D'UN MÉMOIRE DE FIN D'ÉTUDES

AgroParisTech	Mémoire de fin d'études
Titre : Analyse de la performance et des biais liés à la reconnaissance automatisée des chants par BirdNET à partir d'enregistrements acoustiques pour l'étude des communautés d'oiseaux en milieu urbain	Mots clés : Audiomoth, bioacoustique, BirdNET, écologie urbaine, reconnaissance automatisée d'espèces.
Auteur ou autrice : Héloïse Jeux	Année : 2023 - 2024
Caractéristiques : volumes ; pages ; figures ; annexes ; cartes ; bibliographie	

CADRE DU TRAVAIL

Organisme de stage : Laboratoire Image Ville Environnement Nom du responsable : Alejandro Sotillo, François Chiron Fonction : Chercheurs		
Nom de l'enseignant référent d'AgroParisTech : Sébastien Albert		
<input type="checkbox"/> 1A <input type="checkbox"/> 2A <input checked="" type="checkbox"/> 3A	<input type="checkbox"/> Stage entreprise <input type="checkbox"/> Stage assistant ingénieur <input checked="" type="checkbox"/> Stage fin d'études Date de remise :	<input type="checkbox"/> Autre

SUITE À DONNER (à compléter par AgroParisTech)

<input type="checkbox"/> Consultable et diffusable <input type="checkbox"/> Confidentiel de façon permanente <input type="checkbox"/> Confidentiel jusqu'au/...../..... , puis diffusable

Engagement de non-plagiat

❶ Principes

- Le plagiat se définit comme l'action d'un individu qui présente comme sien ce qu'il a pris à autrui.
- Le plagiat de tout ou partie de documents existants constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée.
- Le plagiat concerne entre autres : des phrases, une partie d'un document, des données, des tableaux, des graphiques, des images et des illustrations.
- Le plagiat se situe plus particulièrement à deux niveaux : ne pas citer la provenance du texte que l'on utilise, ce qui revient à le faire passer pour sien de manière passive ; recopier quasi intégralement un texte ou une partie de texte, sans véritable contribution personnelle, même si la source est citée.

❷ Consignes

- Il est rappelé que la rédaction fait partie du travail de création d'un rapport ou d'un mémoire ; en conséquence lorsque l'auteur s'appuie sur un document existant, il ne doit pas recopier les parties l'intéressant mais il doit les synthétiser, les rédiger à sa façon dans son propre texte.
- Vous devez systématiquement et correctement citer les sources des textes, parties de textes, images et autres informations reprises sur d'autres documents, trouvés sur quelque support que ce soit, papier ou numérique en particulier sur internet.
- Vous êtes autorisés à reprendre d'un autre document de courts passages *in extenso*, mais à la stricte condition de les faire figurer entièrement entre guillemets et d'en citer la source.

❸ Sanctions

- En cas de manquement à ces consignes, la direction des études et de la pédagogie ou le correcteur se réservent le droit d'exiger la réécriture du document sans préjuger d'éventuelles sanctions disciplinaires.

❹ Engagement

Je soussignée Héloïse Jeux reconnais avoir lu l'engagement de non-plagiat et je m'engage à le respecter.

À Strasbourg, le 20/09/2024.

Signature :



RESUME :

Ce rapport s'intéresse à l'échantillonnage acoustique des communautés d'oiseaux à l'aide des enregistreurs AudioMoth et de l'algorithme BirdNET, réalisé dans le cadre du projet Ecouter la Ville à Strasbourg. L'objectif était d'évaluer la performance de BirdNET à dresser un relevé correct et complet des communautés présentes le long d'un gradient urbain-forestier soumis à des contextes sonores différents. Les résultats montrent que l'allongement de la durée d'enregistrement améliore la performance de BirdNET si les données sont correctement filtrées, grâce au choix d'un niveau de confiance sévère. D'autre part, l'étude a montré que les caractéristiques des sites urbains et forestiers sont essentielles à la compréhension de la performance de l'algorithme dans les paysages sonores associés, ainsi que les adaptations des espèces s'y trouvant. Cette étude est optimiste sur l'emploi de BirdNET dans des paysages sonores urbains, bien qu'il soit nécessaire d'approfondir les connaissances dans ce domaine.

Mots clés : bioacoustique, écologie urbaine, reconnaissance automatisée d'espèces, Audiomoth, BirdNET.

ABSTRACT :

This report focuses on the acoustic sampling of bird communities using the BirdNET algorithm and AudioMoth recorders. Conducted as part of the Ecouter la Ville project in Strasbourg, France. The objective was to evaluate BirdNET's performance in providing an accurate and comprehensive survey of the communities present along an urban-forest gradient, subject to varying environmental sound conditions. The results show that prolonging the recording duration improves BirdNET's performance when the data are properly filtered by selecting a strict confidence level. Additionally, the study demonstrated that the characteristics of urban and forest sites are crucial to understanding the algorithm's performance in the associated soundscapes, as well as the adaptations of the species found there. This study is optimistic about the use of BirdNET in urban soundscapes, although further research is needed in this field.

Key words : bioacoustics, urban ecology, species automated recognition Audiomoth, BirdNET.

TABLE DES MATIERES

RESUME.....	1
REMERCIEMENTS :	5
LISTE DES FIGURES	6
LISTE DES TABLEAUX	7
ABREVIATIONS.....	8
I) CONTEXTE ET PROBLEMATIQUES.....	9
1. L'échantillonnage acoustique des populations : une méthode avantageuse bénéficiant des avancées récentes en matière d'intelligence artificielle	9
2. L'avifaune : un taxon clé des études en éco-acoustique	10
2.a. Contexte.....	10
2.b. BirdNET & son fonctionnement.....	10
2.c. Des biais à la méthode encore mal connus	11
2.d. Introduction à l'acoustique des oiseaux.....	11
3. Les enjeux liés à l'écologie urbaine	12
3.a. Les villes au cœur d'une situation critique.....	12
3.b. Le projet Evolville : l'écologie urbaine au sein de l'Eurométropole de Strasbourg.....	12
4. Problématique de l'étude & hypothèses	12
4.a. Problématique.....	12
4.b. Hypothèses	13
II) MATERIEL ET METHODES	13
1. Les sites d'études.....	13
2. L'Acquisition des données sur les communautés d'oiseaux grâce à différentes méthodes	14
2.a. Les points d'écoute ou échantillonnage ponctuel	14
2.b. La méthode semi-automatisée :	15
2.b.i. Les AudioMoth : caractéristiques techniques	15
2.b.ii. Pose des Audiomoth sur le terrain	15
2.b.iii. L'analyse des enregistrements par un expert	15
2.c. La méthode automatisée : L'analyse des enregistrements par BirdNET	16
3. L'analyse des performances de BirdNET	16
3.a. La comparaison des méthodes renseigne sur les opportunités et défauts de chacune.	16
3.b. Les performances de BirdNET : la mesure des performances d'un classifieur binaire pour son optimisation	17
3.b.i. Le niveau de confiance comme un filtre de données.....	17
3.b.ii. Des indicateurs de performances multiples permettent d'en cerner les contours	18
3.b.iii. Des outils dédiés au choix d'un modèle : les courbes Precision-Recall.....	18

3.b.iv.	Analyses statistiques générales.....	19
3.b.v.	Comparaisons de communautés grâce aux statistiques	19
4.	Lien entre communauté et paysage	19
5.	Acquisition et traitement des données de pollution sonore	19
6.	Le niveau d'intensité sonore.....	20
6.a.	La nature de la source de son	20
6.b.	Modèles statistiques des effets de la pollution sonore.....	20
6.b.i.	Lien entre pollution sonore et paysage	20
6.b.ii.	Lien entre performance de l'algorithme et pollution sonore	21
III)	RESULTATS	22
1.	Comparaison préliminaire des résultats des méthodes	22
1.a.	Comparaison des communautés obtenues par différentes méthodes.....	22
1.b.	Comparaison des fréquences d'espèces selon les méthodes	23
1.c.	L'hétérogène répartition des niveaux de confiance en fonction des espèces	24
2.	Conséquences de l'allongement de la durée d'enregistrement sur la performance.....	25
2.a.	Effet de la durée d'enregistrement sur les performances de BirdNET	25
2.b.	Effet de la durée d'enregistrement sur le niveau de confiance (i.e. filtre de détections de BirdNET).....	25
3.	Le filtrage des données.....	26
3.a.	Différents filtres correspondant à différents aspects de la performance.....	26
3.b.	Comparaison de ces filtres et des communautés qu'ils produisent	28
4.	Vers des matinées d'enregistrement : le choix d'un filtre	29
4.a.	Validité de la méthode & des filtres pour des durées plus longues	29
4.b.	Comparaison de filtres pour des durées longues d'enregistrement	30
4.c.	Vers le choix d'un filtre	31
5.	Ecologie.....	32
6.	L'étude des biais de la performance de l'algorithme.....	33
6.a.	Caractérisation de la pollution sonore sur les sites.....	33
6.b.	Lien entre les variables de pollution sonore et les caractéristiques paysagères des sites ..	33
6.b.i.	Niveau d'intensité sonore et variables paysagères	33
6.b.ii.	NDSI et variables paysagères	34
6.b.iii.	Taille des sites et gradient urbain	35
6.c.	Lien entre les variables de pollution sonore et performance	35
6.c.i.	Effet du niveau d'intensité sonore sur le niveau de confiance moyen.....	36
6.c.ii.	Effet du NDSI sur le niveau de confiance moyen	37
IV)	DISCUSSION	38
1.	Performance de la méthode automatisée	38

1.a.	Bilan des résultats de la performance de BirdNET	38
1.b.	Les erreurs : les taxons concernés et raisons de ces erreurs	39
1.c.	Importance de la liste d'espèces initiale	39
2.	La pollution sonore est-elle problématique ?	40
2.a.	La taille des sites, une partie de l'explication ?.....	40
2.b.	L'analyse visuelle des spectrogrammes pour différents paysages sonores offre des explications pour le comportement des niveaux de confiance associés aux détections d'espèces	40
2.c.	Récapitulatif et interprétation de l'effet de la pollution sonore sur la performance de BirdNET	42
3.	La question des espèces rares	43
4.	BirdNET : Le rôle des experts naturalistes.....	43
5.	BirdNET : mode d'emploi sur les sites Evolville.....	43
6.	Limites et perspectives de l'étude	44
V)	CONCLUSION	46
	ANNEXE 1 : Liste des espèces fournie à l'algorithme au format BirdNET.	51
	ANNEXE 2 : Niveaux de confiance associés aux détections de BirdNET pour les matinées d'enregistrement.....	53

REMERCIEMENTS :

Je tiens à exprimer ma profonde gratitude à toutes les personnes qui ont contribué à la réalisation de ce mémoire. Ce travail est l'aboutissement d'une période riche en apprentissages et en découvertes et il n'aurait pas été possible sans le soutien et l'accompagnement de nombreuses personnes.

En premier lieu, je remercie chaleureusement mes encadrants, François Chiron et Alejandro Sotillo, pour leur bienveillance et leur pédagogie tout au long de mon stage. Leur expertise et leurs retours constructifs ont été essentiels à l'avancée de mes travaux et à l'élaboration de ce mémoire. Merci pour les journées de terrain, pour m'avoir enseigné les chants et espèces d'oiseaux pour toutes les discussions que nous avons eues.

Je souhaite également adresser mes sincères remerciements aux membres du Laboratoire Image Ville Environnement dans lequel j'ai effectué mon stage. Leur accueil et les échanges enrichissants sur l'écologie urbaine et la recherche m'ont permis de travailler dans un environnement stimulant. Merci pour les discussions sur des thèmes qui nous dépassent et pour les sorties natation.

Je tiens également à remercier mon enseignant référent, Sébastien Albert, pour ses encouragements et son suivi tout au long de ce projet.

Enfin, je n'oublie pas mes proches, à qui je souhaite témoigner toute ma reconnaissance. A ma famille et Leo qui m'aura soutenu dans le *Endspurt*.

Merci à vous.

LISTE DES FIGURES :

Figure 0 : Répartition des sites Evolville sur l'Eurométropole de Strasbourg pour l'année 2024

Figure 1 : Pose des Audiomoths sur le terrain, Avril 2024

Figure 2 : Schéma récapitulatif des 3 méthodes d'échantillonnage acoustique des communautés d'oiseaux.

Figure 3 : Relation entre les données de niveau d'intensité sonore obtenues par des mesures sur le terrain et celles obtenues suite à des calculs automatisés sur R

Figure 4 : Diagramme de Venn comparant les communautés obtenues à l'analyse des mêmes 5 min d'enregistrement, sans filtre sur les données de BirdNET (niveau de confiance égal à 0.1)

Figure 5 : Fréquences des espèces sur l'ensemble du relevé pour chaque méthode et passage

Figure 6 : Niveaux de confiance associées aux points de contacts de chaque espèce du relevé de BirdNET pour 30 min d'enregistrement et un niveau de confiance minimal (0.1)

Figure 7 : Courbe PR figurant la précision en fonction de la sensibilité pour différentes durées d'enregistrement.

Figure 8 : Aire sous les courbes PR en fonction de la durée de l'enregistrement

Figure 9 : Fréquence des catégories de niveaux de confiance associées aux détections pour différentes durées d'enregistrement et proportion d'identifications pour les différentes durées pour chaque catégorie de niveau de confiance

Figure 10 : Valeur des indicateurs de performance en fonction du niveau de confiance sélectionné.

Figure 11 : Courbe PR avec le point correspondant au maximum de la moyenne harmonique de précision et recall.

Figure 12 : Evolution du tableau de contingence pour 30 min d'enregistrement analysées en fonction du niveau de confiance seuil

Figure 13 : Comparaison par NMDS des communautés obtenues avec l'enregistrement et BirdNET pour différentes valeurs de niveau de confiance.

Figure 14 : Evolution des valeurs moyennes des maxima d'indicateurs de performance et des valeurs moyennes de niveaux de confiance associées en fonction de la durée des enregistrements.

Figure 15 : Evolution de la taille du relevé de BirdNET en fonction de différents niveaux de confiance sur les sites de l'étude pour les enregistrements de matinées

Figure 16 : Comparaison des diagrammes de Venn pour NC = 0.69 et NC = 0.95 résumant le nombre d'espèces détectées selon les méthodes.

Figure 17 : Comparaison des communautés obtenues sur les sites grâce à BirdNET sur des enregistrements de matinées en employant deux niveaux de confiance : 0.90 et 0.95.

Figure 18 : Répartition des espèces détectées par BirdNET sur les enregistrements de matinées et des caractéristiques paysagères des sites selon les deux premiers axes de la PCoA.

Figure 19 : Répartition des sites du relevé selon le niveau d'intensité sonore moyen et du NDSI

Figure 20 : Lien entre le niveau d'intensité sonore moyen et NDSI.

Figure 21 : Lien entre le niveau d'intensité sonore moyen et la densité de routes

Figure 22 : Relation entre anthropophonie et biophonie et les éléments du paysage sur les sites

Figure 23 : Comportement du niveau de confiance moyen pour différentes espèces en fonction du niveau d'intensité sonore

Figure 24 : Comportement du niveau de confiance moyen pour différentes espèces en fonction du NDSI

Figure 25 : Spectrogrammes obtenus pour différents sites de l'étude.

Figure 26 : Résumé des étapes clés dans l'utilisation de BirdNET.

LISTE DES TABLEAUX

Tableau 1 : Récapitulatif des durées analysées pour chaque méthode d'échantillonnage acoustique des communautés.

Tableau 2 : Exemple d'une table de contingence obtenue pour analyser les produits d'un classifieur binaire.

Tableau 3 : Description du modèle considérant le niveau de confiance moyen en fonction des 2 variables de pollution sonore et en vérifiant les interactions avec les espèces

Tableau 4 : Description du modèle lm considérant le niveau d'intensité sonore moyen sur 30 min d'enregistrement expliqué par les variables paysagères des sites Evolville

Tableau 5 : Description du modèle lm considérant le NDSI moyen sur 30 min d'enregistrement expliqué par les variables paysagères des sites Evolville

Tableau 6 : Description du modèle lm étudiant le lien entre la taille du site et 2 variables paysagères : la surface végétalisée et la densité de routes dans un rayon de 250 m

Tableau 7 : Description du modèle considérant le niveau de confiance moyen par site, passage et par espèce pour 30 min d'enregistrement en fonction des 2 variables de pollution sonore

ABREVIATIONS :

ACI : Acoustic Complexity Index

ADI : Acoustic Diversity Index

BI : Bioacoustic Index

EMS : Eurométropole de Strasbourg

FN : Faux négatif

FP : Faux positif

FPr : False Positive rate

LIVE : Laboratoire Image Ville Environnement

NC : niveau de confiance

NDSI : Normalized Deviation Standard Index

NDVI : Normalized Difference Vegetalization Index

NMDS : Non Metric Multidimensional Scaling

PCoA : Principal Coordinates Analysis

PR : *Precision-Recall*

ROC : Receiver Operating Characteristic

STOC : Suivi Temporel des Oiseaux Communs

TN : Vrai négatif

TP : Vrai positif

ZAEU : Zone atelier Environnementale urbaine

I) CONTEXTE ET PROBLEMATIQUES

1. L'échantillonnage acoustique des populations : une méthode avantageuse bénéficiant des avancées récentes en matière d'intelligence artificielle

La mesure du vivant est une problématique essentielle dans un contexte de perte drastique de la biodiversité. Les méthodes employées doivent être scrutées avec attention pour en mesurer les avantages mais aussi les inconvénients afin de comprendre l'ampleur des marges d'erreur pour appuyer les décisions politiques et individuelles (Gerber & Iacona, 2024). Les protocoles ont évolué du début de l'histoire des sciences naturelles à nos jours et on observe depuis la deuxième moitié du XXe siècle l'émergence de techniques d'échantillonnage passif, remplaçant la reconnaissance sur le terrain par un expert par l'analyse à partir d'enregistrements visuels ou auditifs, avec notamment l'exemple du piège photographique dès les années 50 (Marcon, 2015).

Dans ce contexte d'innovation, l'écologie acoustique a bénéficié à partir des années 90 de la production d'enregistreurs adaptés à la pose en environnement extérieur et depuis une vingtaine d'années de la mise sur le marché d'enregistreurs programmables dans le temps, tels que le song meter SM4TS (Wildlife acoustics) ou l'AudioMoth (Open Acoustic Devices) (Sugai et al., 2019). Ces enregistreurs offrent de nombreux avantages par rapport à un relevé manuel sur le terrain. En effet, ils permettent d'accéder à des données sur plusieurs heures, voire jours, sur des sites d'intérêt et en simultanément, permettant alors de réduire la contingence dans les observations liée à la courte durée d'une observation directe. Ces enregistrements peuvent ensuite être traités pour analyser les données de différents groupes taxonomiques (Ross et al., 2023). Ainsi, les enregistreurs facilitent le travail de terrain et permettent la réalisation d'études aux saisons froides et/ou de nuit. Il est également possible d'adapter plus facilement l'échelle géographique et temporelle de l'étude aux questions écologiques. Cependant, cette méthode souffre de biais importants intrinsèques et liés à son aspect récent. Par exemple, elle ne permet pas de détecter les espèces silencieuses ou vocalisant peu. L'effet de l'impact du paysage sonore reste à explorer. Elle ne permet pas non plus de s'affranchir totalement des contraintes liées au travail de terrain et au recours d'un expert comme la pose des appareils ainsi que la récupération des données et leur vérification sont nécessaires (Ross et al., 2023). En effet, cette méthode, dont l'intérêt réside dans la quantité de données acquise, nécessite une analyse manuelle qui s'avère chronophage et ne permet de s'affranchir du biais humain que si un seul expert réalise ce travail de vérification (Leach et al., 2016).

Dans l'objectif de valoriser ces grands volumes de données rendus accessibles par les enregistreurs, des algorithmes de reconnaissance automatisée ont vu le jour dans les dernières années, dont on peut citer notamment pour les oiseaux BirdNET (Kahl, 2020), Merlin (Cornell Lab of Ornithology, 2014) et pour les chauves-souris Tadarida (Bas & al., 2017). La diffusion du deep learning et l'avènement des réseaux de neurones à convolution ont amélioré significativement la performance de ces algorithmes qui utilisent dès lors cette technologie (Kahl et al., 2021 ; Xie et al., 2023 ; Napier et al., 2024). Le deep learning se définit comme l'apprentissage automatisé d'un algorithme dans le but d'améliorer ses performances, sans intervention humaine (Journal Officiel de la République Française, 2018). Les réseaux de neurones à convolution appartiennent au deep learning et se caractérisent par une architecture imitant les connexions entre neurones dans un cerveau (Schmidhuber, 2022).

2. L'avifaune : un taxon clé des études en éco-acoustique

2.a. Contexte

Le groupe des oiseaux est considéré comme un indicateur écologique pertinent de l'état des écosystèmes en réponse à certaines pressions anthropiques comme notamment les changements d'usages des sols et de la qualité des habitats (Gregory et al., 2003 ; Lewandowski et al., 2010). En effet, les scientifiques connaissent très bien leur biologie, comportements et traits d'histoire de vie, leur position élevée dans la chaîne trophique renseigne sur des changements à des échelles plus basses (Fraixedas et al., 2020 ; Koskimies, 1989). Le suivi des populations permet d'avoir une estimation de l'impact de changements dans leur habitat à différentes échelles comme ils sont répartis sur l'ensemble du globe et dans différents milieux (Ikin et al., 2016). Par ailleurs, une grande partie des oiseaux émet des vocalisations ou sons, pour attirer un partenaire, marquer un territoire ou autre (Kershenbaum et al., 2016). En conséquence, ce taxon bénéficie des avancées en éco-acoustique et depuis 2014, des chercheurs se rassemblent annuellement lors du BirdCLEF où différents algorithmes de reconnaissance automatisée de chants d'oiseaux sont comparés et leur performance évaluée (Goëau et al., 2014). Des efforts de recherche importants visent à produire des algorithmes performants pour l'identification acoustique des espèces d'oiseaux. Dans ce contexte, il est nécessaire de saisir les biais propres à la combinaison d'enregistrements et à la reconnaissance automatisée. Les recherches à ce sujet sont souvent restées limitées aux milieux naturels (Priyadarshani et al., 2018 ; Bota et al., 2023), en se basant plutôt sur une approche mono-spécifique (Stastny et al., 2018 ; Barré et al., 2019 ; Pérez-Granados, 2023). L'objectif de produire des détections fiables quel que soit l'environnement sonore implique que l'algorithme intègre la complexité et variabilité des chants de chaque espèce (McGinn et al., 2023) et gère les superpositions entre les motifs des vocalisations des espèces d'intérêt et ceux d'autres sons (animaux, vent, pluie, automobiles, etc.) (Xie et al., 2023). Ainsi, les chercheurs conseillent de combiner différentes méthodes d'évaluation de la diversité aviaire pour limiter les retombées des biais d'une méthode sur l'étude bien que les efforts se poursuivent dans le but de produire un algorithme utilisable de façon généralisée à l'échelle des communautés (Besson et al., 2022).

2.b. BirdNET & son fonctionnement

Depuis son apparition, BirdNET s'est imposé comme référence en matière de reconnaissance automatisée (Kahl et al., 2021 ; Pérez-Granados, 2023). Ce réseau de neurones à convolution est issu d'une collaboration entre le Cornell Lab of Ornithology et l'université de Chemnitz. Sa base de données comprend aujourd'hui plus de 6000 espèces (sur environ 11 100 connues sur Terre (Kahl, 2020/2024)). Ce logiciel libre permet d'analyser des enregistrements à travers une interface graphique ou du script Python accessible en ligne, ou bien en direct grâce à une application mobile. Cet algorithme traite un enregistrement en identifiant les motifs du spectrogramme sur des segments de 3s et associe à chaque identification un degré de certitude appelé niveau de confiance (NC) compris entre 0 et 1 (Kahl et al. 2021). Un spectrogramme est un graphique représentant les fréquences des sons obtenus en fonction du temps. La couleur du motif renseigne également sur le niveau d'intensité sonore de ces sons.

Elément clé de l'utilisation de BirdNET, le niveau de confiance permet de filtrer les données de sorties en fonction de sa valeur. Il ne correspond pas à une probabilité et sa formule est :

$$\text{Niveau de confiance} = \frac{1}{1 + e^{\text{logit score} * \text{sensitivity}}}$$

Le *logit score* est le résultat de l'utilisation d'un classificateur linéaire qui intègre des caractéristiques pour produire ce « *score* » non borné. Cette note intègre de multiples aspects influençant la détection comme le contexte sonore, l'espèce, l'appareil de mesure et ses paramétrages, etc. La *sensitivity* ou sensibilité est un paramètre ajustable par l'utilisateur qui modifie le critère d'activation des neurones pour la reconnaissance. Pour résumer, le niveau de confiance est un produit de l'algorithme qui n'est

pas généralisable entre les différentes études (Wood & Kahl, 2024). On peut néanmoins le rapprocher d'un degré de « certitude », comme toutes les études ont montré jusqu'à présent qu'il y a une relation positive entre la précision des détections et ce niveau de confiance (Wood & Kahl, 2024).

2.c. Des biais à la méthode encore mal connus

Trois méthodes d'échantillonnage acoustique de l'avifaune existent dans la littérature : la méthode manuelle correspond aux points d'écoute (ex. Indice Ponctuel d'Abondance de 5, 10 ou 20 min), la méthode semi-automatisée concerne l'enregistrement passif et la reconnaissance auditive des chants à partir de cet enregistrement, et la méthode intégralement automatisée combine enregistrements et reconnaissance par un algorithme, BirdNET dans le cas de cette étude.

En plus de contenir au moins 100 enregistrements par espèce d'oiseaux, la base de données de BirdNET comprend également des enregistrements d'animaux émettant des vocalisations, des sons d'origine géophysique (pluie, orage, etc.), des vocalisations humaines et des sons d'origine anthropique (voitures, avions, sirènes, etc.) et des sons d'enregistreurs acoustiques (Kahl et al. 2021). Bien que BirdNET soit entraîné sur des données potentiellement sources de pollution sonore, les biais qui y sont liés sont mal connus et concernent le paysage (Priyadarshani et al., 2018), le bruit autour du signal rendant son identification par l'algorithme difficile (Priyadarshani et al., 2018), les adaptations des vocalises d'oiseaux à l'environnement sonore (Marín-Gómez & MacGregor-Fors, 2021) et la variabilité individuelle et régionale dans les chants d'oiseaux (Kershenbaum et al., 2016 ; Xie et al., 2023).

Le son autre que les vocalisations d'oiseaux, qui peut être un biais à la reconnaissance, est cependant aussi source d'information sur la nature de l'environnement. En effet, la notion de « paysage sonore » théorisée par Pijanowski, distingue 3 catégories de sons en fonction de leur source. L'anthropophonie correspond aux sons d'origine anthropique, trouvés généralement à des fréquences plus basses (moins de 2-3 kHz), la biophonie regroupe les sons produits par les animaux. Enfin, les sons de source « naturelle » mais non vivante tels que le vent constituent la géophonie (Pijanowski et al., 2011). Un enjeu de cette étude est d'explorer ces deux facettes (importance du biais & description) dans un but méthodologique et de compréhension des paysages. Il existe un enjeu fort autour de la compréhension de l'effet des composantes du paysage sonore sur les détections de BirdNET, notamment en milieu urbain où l'emploi de cette technologie reste limité.

2.d. Introduction à l'acoustique des oiseaux

Les oiseaux émettent des sons dans une gamme de fréquences généralement comprise entre 1 kHz et 8 kHz (Brumm & Slabbekoorn, 2005) qui vont de chants mélodieux comme la grive musicienne (*Turdus philomelos*) à des cris comme le corbeau (*Corvus corax*).

L'environnement influence les chants des oiseaux. En forêt, où les arbres et le feuillage dense peuvent étouffer les sons aigus, certaines espèces adoptent des fréquences de chant plus basses afin de favoriser leur propagation (Morton, 1975). Dans les environnements urbains, où la circulation notamment est source de pollution sonore, l'adaptation des espèces est critique et se caractérise par leur plasticité acoustique, i.e. leur faculté à modifier leur comportement de chant (Slabbekoorn & Peet, 2003). Par exemple, elles augmentent la fréquence de leurs vocalises pour limiter les overlap avec l'anthropophonie caractérisée par des fréquences basses (1 à 3kHz) (Slabbekoorn & Peet, 2003). Des changements de phénologie temporelle ont également été observés avec des espèces commençant à chanter plus tôt le matin avant que l'anthropophonie ne devienne trop puissante (Nordt & Klenke, 2013). Les espèces chantent également plus, plus lentement et ont des chants plus longs que les espèces non-urbaines (Marín-Gómez & MacGregor-Fors, 2021).

3. Les enjeux liés à l'écologie urbaine

3.a. Les villes au cœur d'une situation critique

L'écologie urbaine est une discipline scientifique ayant pris son essor à la fin du XX^e siècle (Niemelä, 2011). Les milieux urbains sont définis par une densité humaine et de bâti importante et incluent les zones industrielles et les voies de communication (Muratet, 2006). Depuis 2007, plus de la moitié de la population mondiale est urbaine (UNDESA, 2016) et le changement climatique remet en question l'organisation des villes dont les matériaux de construction et la géométrie sont sources de maux grandissants : îlot de chaleur urbain, imperméabilité des sols, pollution etc. (*UN News*, 2014). L'arrivée de l'écologie urbaine signe un changement de paradigme pour lequel les villes sont à présent considérées comme des écosystèmes. Ainsi, de multiples études témoignent de la nécessité de repenser l'espace urbain en incluant la biodiversité comme celle-ci fait partie de la solution pour la résilience des villes (Muratet, 2006).

3.b. Le projet Evolville : l'écologie urbaine au sein de l'Eurométropole de Strasbourg

La conservation des espaces de nature en ville est un enjeu critique. Ces espaces sont souvent soumis à une pression foncière forte et souvent ne bénéficient pas d'une protection légale. Par exemple, les friches ne sont pas considérées comme des « espaces verts », ce qui résulte dans le cadre de la loi Zéro Artificialisation Nette de 2023, en leur perception comme d'une « richesse à exploiter » (GOIDIN et al., 2022). Or ces espaces non gérés sont d'excellents refuges de biodiversité en ville (Bonthoux et al., 201 ; Muratet et al., 2021). Le projet Evolville, mené par le Laboratoire Image Ville Environnement est un suivi, initié en 2020, de 60 friches, pelouses et gazons au sein de l'Eurométropole de Strasbourg. Il vise à mesurer la biodiversité le long de gradients urbains, pour comprendre les interactions entre les communautés biologiques, le paysage et la gestion des sites et appuyer la valeur écologique de ces sites par un suivi scientifique. Un suivi floristique est réalisé annuellement depuis 2021 et les variables paysagères et de gestion des sites sont connues (Muratet et al., 2024).

4. Problématique de l'étude & hypothèses

4.a. Problématique

L'étude des communautés d'oiseaux comme indicateur de la qualité des écosystèmes urbains s'impose au regard des explications précédentes dans le suivi des sites Evolville. Le stage s'intègre dans cette démarche et s'intéresse à la méthodologie de ce suivi. La problématique guidant l'étude est la suivante : les performances de la méthode automatisée, combinant l'usage d'enregistreurs Audiomoth et de l'algorithme de classification automatique BirdNET, le long du gradient urbain-forestier permettent-elles son utilisation selon une méthodologie unifiée pour reproduire une image fiable des communautés d'oiseaux ?

L'étude se place à l'échelle des communautés, où l'enjeu est l'enregistrement de la présence ou absence de l'espèce sur un site. Plus commune dans la littérature, l'approche spécifique regarde les performances de BirdNET dans la détection et reconnaissance de chaque motif à l'échelle d'un spectrogramme. L'originalité de cette étude réside dans sa considération d'un gradient incluant des environnements acoustiques urbains et s'intéressant à son effet sur la performance de la méthode automatisée. Cette étude se veut précurseur dans le cadre du projet *écouter la ville* mené par le LIVE, dont l'enjeu, à terme, sera d'intégrer le suivi acoustique des chiroptères, oiseaux et éventuellement d'autres taxons comme les orthoptères.

4.b. Hypothèses

Les hypothèses de l'étude sont les suivantes :

- Utiliser des durées longues d'enregistrement permet d'améliorer la performance de la méthode automatisée.
- L'aspect de la pollution sonore nuisant à la reconnaissance automatisée d'espèces est le niveau d'intensité sonore. La nature du son n'a pas d'effet.
- Il existe un gradient anthropophonie dominante – biophonie dominante dans le paysage sonore des sites d'étude qui reflète la répartition des communautés d'oiseaux.

II) MATERIEL ET METHODES

1. Les sites d'études

L'étude comprend 60 sites d'étude répartis sur l'ensemble de l'Eurométropole de Strasbourg (Bas-Rhin, 339,85 km² (*Strasbourg.eu*)) (figure 0) correspondant à des milieux herbacés mésophiles ouverts. Ces sites sont répartis le long d'un gradient allant de sites urbains à des sites forestiers. Trois catégories de milieux ouverts sont suivies, que l'on peut distinguer par leur degré d'entretien. Les gazons sont maintenus ras grâce une fauche fréquente (12-16 fois par an). Ils sont trouvés le plus souvent dans des parcs ou des endroits très fréquentés. Les prairies sont fauchées moins souvent, ce qui permet de laisser monter les plantes en graines notamment et d'avoir une végétation herbacée plus haute. Enfin, les friches sont des espaces ouverts peu ou pas gérés, souvent issus de l'arrêt des activités humaines sur un espace, dans l'attente d'une nouvelle occupation (*Journal officiel*, 1998).

Ces sites se trouvent dans le Bas-Rhin, où le climat est tempéré et semi-continental. Le projet est porté par le Laboratoire image, ville, environnement de l'Université de Strasbourg, en collaboration avec l'Eurométropole de Strasbourg (EMS), le jardin botanique de l'Université de Strasbourg et la Zone atelier environnementale urbaine (ZAEU).

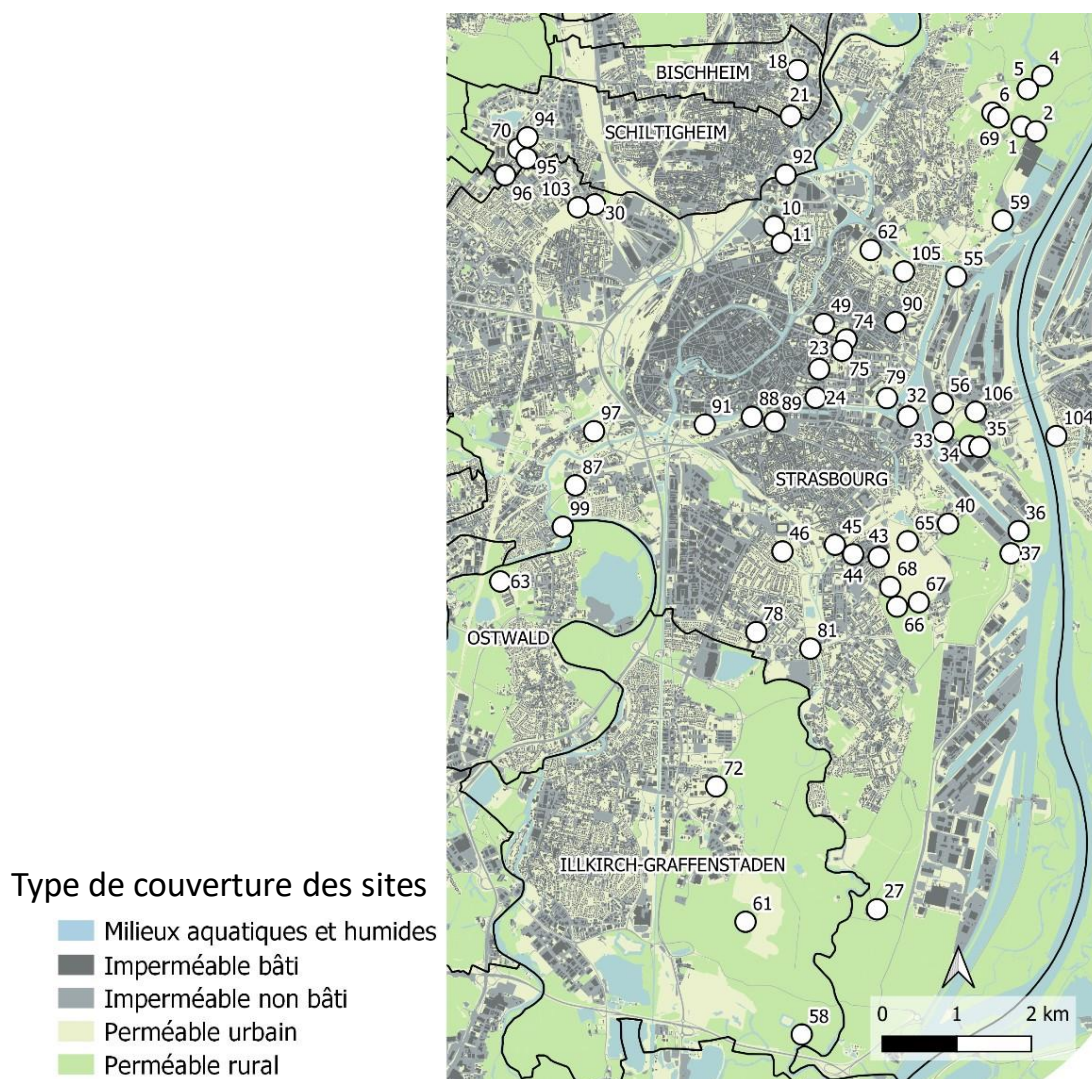


Figure 0 : Répartition des sites Evolville sur l’Eurométropole de Strasbourg pour l’année 2024

2. L’Acquisition des données sur les communautés d’oiseaux grâce à différentes méthodes

2.a. Les points d’écoute ou échantillonnage ponctuel

Les communautés d’oiseaux ont été échantillonnées à partir de points d’écoute (un par site) réalisés pendant la période de reproduction par un ornithologue compétent (FC). Le relevé suit le protocole du Suivi Temporel des Oiseaux Communs (STOC, VigieNature) qui comprend 2 passages par site au printemps. Le premier passage s’effectue en avril, afin de détecter les premiers reproducteurs, et le second 4 à 6 semaines après (mai-juin), suivant l’arrivée et l’installation des derniers migrateurs. Au cours d’un passage, toutes les espèces et individus vus et entendus pendant 5 minutes et dans trois rayons de distance autour de l’observateur (à moins de 25 m, entre 25 et 100 m, plus de 100 m), sont notés. Le point d’écoute a lieu aux horaires d’activité maximale des oiseaux, entre le lever du soleil et 9h, c’est-à-dire juste après le chorus matinal ainsi limiter l’effet de la phénologie temporelle sur le relevé (STOC, VigieNature).

2.b. La méthode semi-automatisée :

2.b.i. Les AudioMoth : caractéristiques techniques

Les AudioMoth sont des enregistreurs de petite taille issus du travail de chercheurs et d'ingénieurs de l'université de Southampton et aujourd'hui produits par l'entreprise Open Acoustic Devices. Leur mise sur le marché en 2018 a contribué à la diffusion des techniques d'écoacoustique en écologie comme ils sont économes en énergie, produisent des enregistrements d'une qualité suffisante pour réaliser des études et de faible coût (80 USD) (Hill et al., 2018 ; Bota et al., 2023). Ces appareils enregistrent des durées de 5 min en continu sur les horaires programmés.

2.b.ii. Pose des Audiomoth sur le terrain

Chaque passage correspond à l'étude de 60 sites et donne lieu à la pose d'une vingtaine d'enregistreurs sur autant de sites. Ainsi 3 phases de terrain sont nécessaires pour un passage. L'enregistrement commence en début de nuit et se termine vers 9 h pour capturer l'activité nocturne des chiroptères utile à d'autres études et l'activité matinale des oiseaux, sur une période concomitante avec le point d'écoute. Deux nuits consécutives sont enregistrées sur chaque site. Les nuits sont choisies de manière à ce que la température ne descende pas en dessous de 11°C, afin que l'activité des insectes et donc des chiroptères soit suffisante pour la détection. Après un formatage initial des AudioMoth, ils sont enfermés dans des capsules étanches, avec une ouverture protégée au niveau du micro. Ils sont alors fixés hors d'atteinte sur des branches d'arbres en lisière de site. Après 2 nuits d'enregistrement, ils sont récupérés, les enregistrements et les caractéristiques de l'appareil sont récupérés.



Figure 1 : Pose des Audiomoth sur le terrain, Avril 2024 - François Chiron et Héroïse Jeux - Photo d'Alejandro Sotillo

2.b.iii. L'analyse des enregistrements par un expert

Sur les 60 sites Evolville, le suivi acoustique à l'aide des enregistreurs Audiomoth a permis de retenir 31 sites (29 communs aux deux passages et 2 avec un seul passage) pour lesquels les conditions d'écoute (météorologie clémente, absence de vent et pluie) et la réalisation (enregistrement complet) ont été satisfaisantes. Deux passages ont été réalisés en 2024 et sont considérés distinctement comme les communautés observées sur les sites sont différentes pour des raisons de phénologie temporelle.

Après avoir rassemblé tous les enregistrements pour chaque passage, les séquences d'enregistrements correspondant à la plage horaire des points d'écoute sur le terrain ont été extraites, soit entre 6h30 et 9h pour les enregistrements du mois d'avril et entre 5h30 et 9h pour ceux du mois de juin. L'ensemble de ces enregistrements matinaux vont servir de support à l'analyse de BirdNET. Seul un échantillon de ces « matinées » va être analysée par un expert. Ainsi, après vérification que les communautés d'oiseaux de 2 matinées consécutives d'enregistrement ne sont pas significativement différentes grâce à une Permanova (fonction Adonis2 et une NMDS), 6 séquences de 5 min d'enregistrement ont été échantillonnées pour chaque site et passage. Ces séquences correspondent à 3 enregistrements par matinée, choisis à des horaires similaires d'une matinée à l'autre afin de s'affranchir du biais horaire, mais espacés de 45 minutes pour une même matinée. Parmi ces 6 fois 5min d'enregistrement, 5 minutes sont concomitantes au point d'écoute du site. Ces enregistrements sont écoutés par un expert et les espèces entendues relevées.

2.c. La méthode automatisée : L'analyse des enregistrements par BirdNET

Pour cette étude, la version V2.4 de BirdNET a été utilisée. Elle peut détecter plus de 6000 espèces d'oiseaux réparties à l'échelle mondiale et couvre des fréquences comprises entre 0 et 15 kHz (Kahl, 2020). L'application BirdNET a été paramétrée à partir de son interface graphique. Les valeurs de « *sensitivity* » et « *overlap* » ont été choisies grâce à l'étude de D. Funosas sur l'effet de leur valeur sur la performance de l'algorithme. La « *sensitivity* » qui intervient dans la production du score de confiance associé à une détection a été fixée à 1.5 (Funosas et al., 2024). L'« *overlap* » est une donnée comprise entre 1 et 3 secondes. Il permet lors de l'analyse d'un segment de 3 secondes que l'algorithme étende ce segment à chacune de ses extrémités d'une durée égale à l'overlap pour faciliter la reconnaissance de motifs vocaux qui seraient coupés. Pour cette étude, il est donc fixé à 2s (Funosas et al., 2024). Une entrée importante pour le fonctionnement de l'algorithme consiste à choisir la liste d'espèces de référence. La liste employée est issue de la liste des espèces observées à l'échelle de l'Alsace sur la période 2015-2024 (de février à juillet) fournie par le site '*faune-alsace*'. Elle est raccourcie des espèces non-nicheuses et des espèces dont la présence est impossible au sein de l'EMS (espèces typiques des milieux montagnards par exemple (Grand Tétrás), etc.). Au total elle compte 144 espèces. Enfin, le niveau de confiance minimal pour sauvegarder une détection de l'algorithme est fixé à la valeur de 0.1 que l'on retrouve dans la littérature (Funosas et al. 2024, Bota et al., 2023).

3. L'analyse des performances de BirdNET

3.a. La comparaison des méthodes renseigne sur les opportunités et défauts de chacune.

L'analyse des performances de la méthode automatisée implique la mesure et la comparaison du signal à 3 étapes : le point d'écoute, l'analyse de l'enregistrement par un expert, et l'analyse de l'enregistrement par BirdNET (Figure 2). La comparaison des identifications entre l'écoute des enregistrements et le point d'écoute nous renseigne sur ce qui est perdu lors de l'enregistrement, soit à cause de la sensibilité du microphone, soit parce que les espèces sont seulement vues. La comparaison entre l'écoute de l'enregistrement et les sorties de BirdNET nous permet de quantifier la performance de l'algorithme. Cette comparaison permet d'optimiser les paramètres de BirdNET afin d'obtenir la meilleure performance possible.

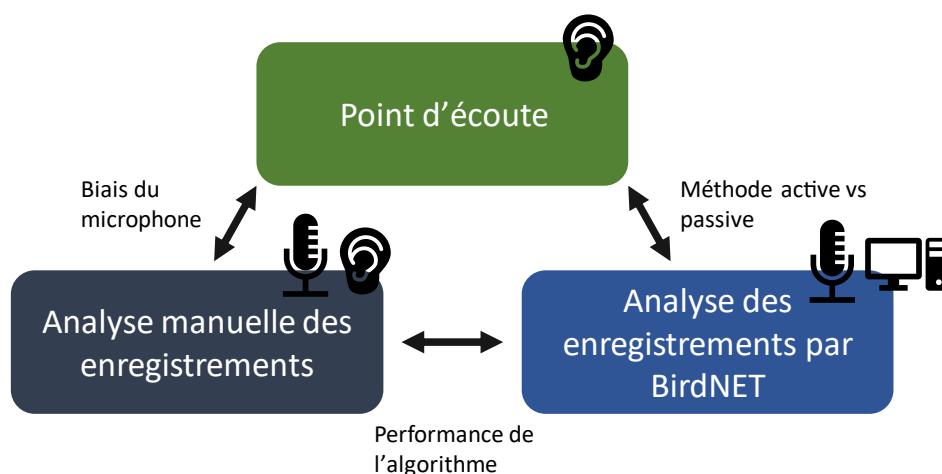


Figure 2 : Schéma récapitulatif des 3 méthodes d'échantillonnage acoustique des communautés d'oiseaux. Les flèches représentent les enjeux principaux liés à la comparaison de 2 méthodes.

L'avantage principal de la méthode automatisée est le traitement automatisé de longues durées d'enregistrement sur chaque site, qui est un enjeu de l'étude. Ainsi, l'étude s'applique également à comparer les informations obtenues sur les communautés des sites pour différentes durées et méthodes (Tableau 1).

Durée analysée	Comparaison des méthodes de détection des espèces		
	Point d'écoute	Analyse de l'enregistrement manuelle	BirdNET
5 min	X	X	X
30 min (6 fois 5min)		X	X
“matinée”			X

Tableau 1 : Récapitulatif des durées analysées pour chaque méthode d'échantillonnage acoustique des communautés.

3.b. Les performances de BirdNET : la mesure des performances d'un classifieur binaire pour son optimisation

3.b.i. Le niveau de confiance comme un filtre de données

BirdNET est un exemple de classifieur binaire. Son mode de fonctionnement correspond à la classification d'un objet soit dans la catégorie « évènement » ou 1, soit dans la catégorie « non-évènement » ou 0. Dans l'étude, cet objet se rapporte aux segments et motifs du spectrogramme analysés. L'« évènement » va être la détection d'une espèce, le « non-évènement » l'absence de détection. Il s'agit alors de classer ces détections et non-détections pour identifier les erreurs par rapport à la référence. Dans le cas de l'étude des performances de l'algorithme, il s'agit des données issues de l'analyse de l'enregistrement par l'expert. On peut alors distinguer 4 catégories : les vrais positifs et vrais négatifs sont les réponses correctes, tandis que les faux négatifs et faux positifs sont les erreurs (Tableau 2).

Identifications	Entendu par l'expert	Pas entendu par l'expert
Validé par la méthode testée	Vrais positifs (TP)	Faux positif (FP)
Ignoré par la méthode testée	Faux négatifs (FN)	Vrais négatifs (TN)

Tableau 2 : Exemple d'une table de contingence obtenue pour analyser les produits d'un classifieur binaire. Le vert souligne les réponses correctes, tandis qu'en rouge se trouvent les erreurs.

L'étude est réalisée à l'échelle des communautés pour chaque site. Ainsi, un vrai positif (True Positive TP) correspond à une espèce relevée par les deux méthodes. Un vrai négatif (True Negative TN) correspond à une espèce qui n'est pas sur le site d'après l'expert et n'est pas entendu par la méthode test. Un faux positif (False Positive FP) est une erreur de la méthode test qui renvoie une espèce qui n'est pas sur le site, alors qu'un faux négatif (False Negative FN) n'est pas détecté par la méthode test alors qu'elle est relevée par l'expert.

Le choix d'un niveau de confiance lors du paramétrage de BirdNET par l'utilisateur définit un filtre. Pour les valeurs de niveau de confiance inférieures à ce filtre, les détections seront ignorées. Ainsi, l'optimisation du taux d'erreur se fait à travers le choix d'un niveau de confiance seuil pour l'ensemble des sites et passages qui va permettre de moduler le compromis (ou 'trade-off') entre sensibilité et véracité de la détection. Ainsi, des seuils de niveaux de confiance ont donc été construits, compris dans l'intervalle [0.1, 1] avec un pas de 0.01 et les analyses de BirdNET étudiées.

3.b.ii. Des indicateurs de performances multiples permettent d'en cerner les contours

Pour chacun de ces seuils, des indicateurs permettant de quantifier les différents aspects de la performance sont calculés pour chaque site et passage :

- Precision = $TP / (TP + FP)$: part de prédictions correctes sur l'ensemble des espèces détectées par site, nommée aussi précision.
- Recall = $TP / (TP + FN)$: sensibilité de l'algorithme à la détection des espèces
- Le F β -score correspondant à la moyenne harmonique de la *Precision* et du *Recall* (Flach & Kull, 2015) (moyenne adaptée aux pourcentages) permet de quantifier le *trade-off* entre ces deux indicateurs. Le facteur β permet de donner plus de poids à la précision si $\beta < 1$, de favoriser le *recall* si $\beta > 1$ et d'équilibrer *precision* et *recall* si $\beta = 1$.

$$F\beta_{score} = \frac{(\beta^2 + 1) \cdot precision \cdot recall}{\beta^2 precision + recall}$$

En postulant que la précision a plus d'importance que la sensibilité dans le cadre de suivi à l'échelle des communautés (Funosas et al., 2024), ont été considérées les valeurs de β suivantes : 1, 0.75, 0.5, 0.25.

- FPR = $FP / (FP + TN)$: part de détections fausses parmi ce qui n'est pas à détecter
- Similarité de Jaccard : indice de similarité entre deux communautés

$J = c / (a + b - c)$, avec a et b les nombres d'espèces des communautés A et B, c le nombre d'espèces communes entre ces 2 communautés

Ces indicateurs, calculés pour différents niveaux de confiance et durée, donnent accès à un des niveaux de confiance optimaux pour différents aspects de la performance.

3.b.iii. Des outils dédiés au choix d'un modèle : les courbes Precision-Recall

Des outils de *machine-learning* permettent de capturer ces enjeux et de choisir un modèle. Les courbes ROC (*Receiver Operating Characteristic*) et courbes PR (*Precision-Recall*) sont des méthodes de représentation graphique et de calcul de la performance d'un modèle de classification binaire. Ces courbes permettent de comparer notre classifieur à un modèle de classification aléatoire grâce à l'AUC (Area Under the Curve – aire sous la courbe) pour lequel cette mesure sera 0.5 (proportion aléatoire d'éléments correctement classés par un modèle aléatoire) tandis qu'elle vaut 1 pour un modèle parfait.

La courbe ROC trace le Recall (taux de TP) en fonction du FPr (Taux de FP), chaque point correspondant à un niveau de confiance dans notre étude. Ainsi, cette courbe renseigne sur le niveau de confiance optimal à utiliser correspondant au niveau de confiance associé au point de la courbe le plus proche du point de coordonnées (0,1). L'intérêt de cette courbe réside dans le fait qu'elle montre le *trade-off* entre la sensibilité de la détection et le taux de FP qui est l'erreur que l'on cherche à éviter. L'AUC de la courbe ROC correspond au pourcentage de positifs réels correctement classés comme positifs par le classifieur.

La courbe PR (*Precision-Recall*) ($Precision = f(Recall)$) figure le *trade-off* entre la proportion de réponses correctes parmi l'ensemble des détections et la sensibilité de la détection. Le point optimal est à proximité du point de coordonnées (1,1), renseignant alors également sur le niveau de confiance à employer. Un consensus existe dans la littérature sur l'emploi privilégié d'une ou l'autre des 2 courbes.

En effet, la courbe ROC tient compte dans sa construction des vrais négatifs (TN). Or dans le cas où les classes positives et négatives sont très déséquilibrées à cause du grand nombre de TN, l'usage de cette courbe n'est pas recommandé, puisque les variations de FP dans $FPr = FP / (FP + TN)$ seront gommées menant à une vision optimiste des résultats de l'algorithme (Davis & Goadrich, 2006). C'est le cas de cette étude où le nombre de TN excède largement les autres catégories, causant un déséquilibre et poussant à l'utilisation préférentielle de la courbe PR.

L'utilisation de l'AUC de la courbe PR prête cependant à controverse comme il n'existe pas d'interprétation de cette surface (Flach & Kull, 2015). L'utilisation de cette méthode consiste alors à comparer la position des courbes PR pour différents modèles et en déduire le meilleur (courbe la plus proche du modèle idéal = courbe la plus "au-dessus") (Davis & Goadrich, 2006). La courbe PR renseigne également sur le niveau de confiance à utiliser. Il s'agit du niveau de confiance associé au point de la courbe le plus proche du coin haut-droit, de coordonnée (1,1). Ce point de la courbe PR correspond alors au niveau de confiance maximisant la moyenne harmonique de Precision et Recall, i.e. le niveau de confiance pour lequel F1-score est maximal (Davis & Goadrich, 2006).

3.b.iv. Analyses statistiques générales

Les packages généraux utilisés dans cette étude sur R sont : dplyr (Whickham et al., 2023), tidyr (Whickham et al., 2023) & ggplot2 (Whickam, 2016) pour les graphes, stats (R Core Team, 2023) pour les analyses statistiques de base, Seewave (Sueur & al., 2018) pour la gestion des fichiers au format .wav. Les analyses et calculs décrits par la suite ont été réalisés avec R version 4.4.0 (2024-04-24 ucrt).

3.b.v. Comparaisons de communautés grâce aux statistiques

Pour comparer les compositions des communautés obtenues selon différentes méthodes, durées d'enregistrement ou niveau de confiance utilisé, des NMDS (Non-Metric Multidimensional Scaling) ont été réalisées à l'aide du package vegan (Oksanen et al., 2022) sur les matrices de distance de Jaccard. Les variations entre ces communautés ont été testées à l'aide de la fonction adonis2 (vegan) correspondant à une permanova (analyse de variance par permutations) et de la fonction Anosim (vegan) pour des tests de rangs basés sur la similarité. Les Anosim produisent une valeur R correspondant à la différence entre la moyenne des rangs inter et intra-groupes, comprise entre -1 (différences entre les rangs intra-groupes plus importantes) et 1 (différences entre les rangs inter-groupes dominant)(Somerfield et al., 2021).

4. Lien entre communauté et paysage

Dans le but d'évaluer la méthode de détection automatisée des communautés d'oiseaux afin de pouvoir étudier leur lien avec les paramètres de paysage, il est primordial de vérifier si la méthode permet de refléter le gradient urbain-forêt le long duquel est construit le choix des sites d'étude. La PCoA (*Principal Coordinates Analysis*) (package stats) nous permet de figurer la répartition des espèces selon leur spécialisation écologique à partir d'une matrice de distance de Jaccard.

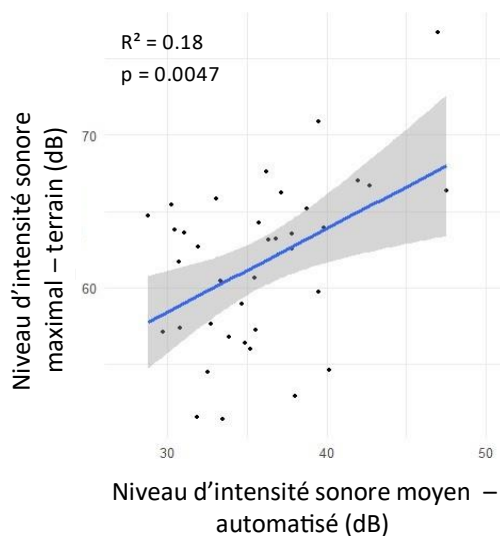
5. Acquisition et traitement des données de pollution sonore

Une multitude d'indicateurs existent pour caractériser un paysage sonore à travers un spectrogramme (Pieretti et al., 2011 ; Villanueva-Rivera et al., 2011 ; Boelman et al., 2007 ; Kasten et al., 2012). Il est cependant nécessaire d'employer une combinaison de ces indicateurs comme aucun ne caractérise à lui seul la complexité acoustique d'un environnement (Bradfer-Lawrence et al., 2023). De plus, étant donné les interactions entre les indicateurs qui existent, l'objet de leur caractérisation et l'influence du contexte sonore sur cette caractérisation sont encore mal connus, notamment en milieu urbain, l'emploi de ces indicateurs est sujet à précautions (Fairbrass et al., 2017). Ainsi, le choix des indicateurs de l'étude s'est

porté sur la faisabilité de leur mesure et la pertinence de leurs résultats. Deux variables de son ont été retenues pour l'étude.

6. Le niveau d'intensité sonore

Le niveau d'intensité sonore (L) correspond au volume sonore et se calcule à partir de l'intensité sonore I (W.m^{-2}) et du seuil d'audibilité I_0 constant (W.m^{-2}) : $L = 10 \log (I/I_0)$. Lors de chaque point d'écoute sur le terrain, les valeurs maximales de niveau d'intensité sonore ont été mesurées à l'aide d'un sonomètre (Digital Sound Level Meter GM1351, Benetech®) en s'orientant vers chaque point cardinal successivement pendant 1 min ($t=4$ minutes au total). D'autre part, le package tuneR (Ligges & al., 2023) a permis de calculer le niveau d'intensité sonore moyen et maximal pour les 6 x 5 min d'enregistrement par site et par passage.



La relation entre les données de niveau d'intensité sonore maximal relevées sur le terrain et le niveau d'intensité sonore moyen produit par R à partir des enregistrements des Audiomoth autorise l'utilisation des données produites par R, permettant d'augmenter le jeu de données de pollution sonore et de calculer niveau d'intensité sonore et NDSI sur 30 min d'enregistrement par site et passage (figure 3).

Figure 3 : Relation entre les données de niveau d'intensité sonore obtenues par des mesures sur le terrain et celles obtenues suite à des calculs automatisés sur R. Les points représentent les valeurs obtenues pour 5 min

6.a. La nature de la source de son

L'étude s'est également intéressée à la nature du son, en se concentrant sur la distinction entre son biologique et anthropique. Le Normalized Deviation Standard Index (NDSI) calcule un ratio entre anthropophonie et biophonie pour un enregistrement (Kasten & al., 2012). Cet indicateur est borné entre -1 et 1, les valeurs négatives reflétant les paysages sonores où le bruit anthropique domine et les valeurs positives les paysages sonores où la biophonie domine.

Sa formule s'écrit $\text{NDSI} = (\beta - \alpha) / (\beta + \alpha)$, avec α et β la densité spectrale de puissance maximale pour des segments de spectre de 1 kHz correspondant respectivement à l'anthropophonie et à la biophonie (Kasten & al., 2012). Le calcul des valeurs de l'indicateur sur R s'est fait sur les 6 x 5min d'enregistrement pour chaque site par passage, en considérant que la biophonie concernait les fréquences entre 2 et 8kHz et l'anthropophonie les fréquences entre 1 et 2 kHz (Kasten & al., 2012). Le package Soundecology (Villanueva-Rivera & Pijanowski, 2018) permet de calculer le NDSI sur R.

6.b. Modèles statistiques des effets de la pollution sonore

6.b.i. Lien entre pollution sonore et paysage

Pour aborder la question de la pollution sonore le long du gradient urbain-forestier, une première étape a été de chercher à comprendre le lien entre le niveau d'intensité sonore et le NDSI (variables dépendantes) et le paysage sur les sites Evolville (variables explicatives). Un modèle (lm) par variable de pollution sonore a été fait et des anovas de type II ont permis de mesurer le pouvoir explicatif de chaque variable paysagère. Les variables de pollutions sonores correspondent à des moyennes par site

pour chaque passage sur 30 min d'enregistrement. Les variables paysagères ont été sélectionnées de manière à ce qu'elles ne soient pas corrélées et comprennent :

- Le Normalized Difference Vegetation Index (NDVI) : indice de végétalisation des sites. On considère le NDVI dans un rayon de 250 m à partir du centre du site.
- La fréquence de fauche par an
- La compacité du bâti dans un rayon de 300 m
- La densité de population dans un rayon de 150 m
- La densité de routes dans un rayon de 150 m

Les variables paysagères et de gestion des sites d'étude sont connues et issues de phases de terrain depuis 2020 et des questionnaires des sites.

6.b.ii. Lien entre performance de l'algorithme et pollution sonore

Dans un second temps, le lien entre la performance de BirdNET et la pollution sonore a été étudiée. Le niveau de confiance moyen sur 30 min, par site, pour chaque espèce observée par la référence sur le site (i.e. seuls les vrais positifs sont considérés) a permis de caractériser cette performance. Ce choix s'explique par la volonté d'avoir un indicateur pertinent n'étant pas totalement soumis à la contingence du point de contact, qui peut avoir lieu lorsque l'oiseau est sur la même branche que l'Audiomoth ou à l'opposé sur le site d'étude. De plus, ceci permet d'avoir une information "brute" de la performance de l'algorithme et donc une mesure de performance de l'algorithme plus intuitive et informative. Les espèces concernées sont communes sur les sites de l'étude : Corneille noire, Chardonneret élégant, Fauvette à tête noire, Grimpereau des jardins, Merle noir, Mésange bleue, Mésange charbonnière, Moineau domestique, Pic épeiche, Pic vert, Pigeon ramier, Pinson des arbres. D'autres sont moins fréquentes : Corbeau freux, Faucon crécerelle, Lorient d'Europe, Mésange nonnette, Pouillot véloce, Rougequeue noir, Troglodyte mignon, Verdier d'Europe. Ceci permet d'avoir un aperçu de l'effet de la pollution sonore sur ces deux catégories d'espèces

Un premier modèle (beta-régression) a donc considéré le lien entre niveau de confiance moyen (variable dépendante) et les variables de pollution sonore (variables explicatives), ainsi que l'interaction entre espèces et les variables de pollution sonore comme on s'attend que l'effet de la pollution sonore ne soit pas le même sur toutes les espèces (Tableau 3). Une anova de type III prenant en compte les interactions permet de vérifier le pourcentage de variance expliqué par chaque variables et les interactions.

Variables	Valeur de la statistique (Chisq)	p-value	Df
Intercept	1.96	0.16	1
Niveau d'intensité sonore	0.95	0.33	1
NDSI	1.29	0.26	1
Interaction niveau d'intensité sonore et espèce	76.1	1.77e-8	20
Interaction NDSI et espèce	41.9	2.8e-3	20

Tableau 3 : Description du modèle considérant le niveau de confiance moyen en fonction des 2 variables de pollution sonore et en vérifiant les interactions avec les espèces. Les espèces considérées sont des vrais positifs, i.e. elles ont été entendues par l'expert sur les sites en question. Les données sont calculées par site et passage pour 30 min d'enregistrement.

Comme cette beta régression montre que l'effet de la pollution sonore est spécifique (tableau 3), on un modèle mixte a été testé pour les espèces présentes sur les 2 passages, avec le passage en effet aléatoire, qui s'est avéré ne pas être significatif. Le package glmmTMB a été utilisé pour construire des modèles

mixtes afin d'étudier l'effet de la pollution sonore, en appliquant une régression bêta comme famille de distribution. Le package effects a été employé pour explorer et visualiser les effets des variables du modèle. On a donc utilisé une beta régression pour les modèles spécifiques dans le but d'étudier l'effet des variables NDSI et niveau d'intensité sonore moyenne (variables explicatives) sur ce niveau de confiance moyen (variable dépendante) par espèce. La valeur de pseudo R^2 s'applique aux données de niveau de confiance qui ne respectent pas l'hypothèse de normalité et permet de mesurer l'ajustement du modèle. Cette étude utilise le pseudo R^2 de Nagelkerke qui compare les vraisemblances du modèle ajusté et du modèle sans prédicteurs en tenant compte du nombre d'observations. Ce pseudo- R^2 est borné entre 0 et 1 (Cribari-Neto & Zeileis, 2010).

III) RESULTATS

1. Comparaison préliminaire des résultats des méthodes

Une analyse des résultats préliminaires des analyses renseigne sur les enjeux liés aux méthodes.

1.a. Comparaison des communautés obtenues par différentes méthodes

Une comparaison simple des communautés obtenues sur la même durée avec la méthode manuelle (point d'écoute), semi-automatisée et automatisée (BirdNET) s'impose. En passant de la méthode manuelle à semi-automatisée, on observe une perte de 13 espèces, qui s'explique par le fait que certaines sont des espèces vues et non entendues (grand cormoran (*Phalacrocorax carbo*), canard colvert (*Anas platyrhynchos*), Goéland Argenté (*Larus argentatus*)), ou alors elles sont plus difficiles à détecter comme elles émettent des cris (Faucons, buse variable (*Buteo buteo*)) (figure 4). Enfin, la perte des espèces restantes peut s'expliquer par leur faible présence sur les sites, rendant moins probable leur détection à l'échelle de tout le relevé.

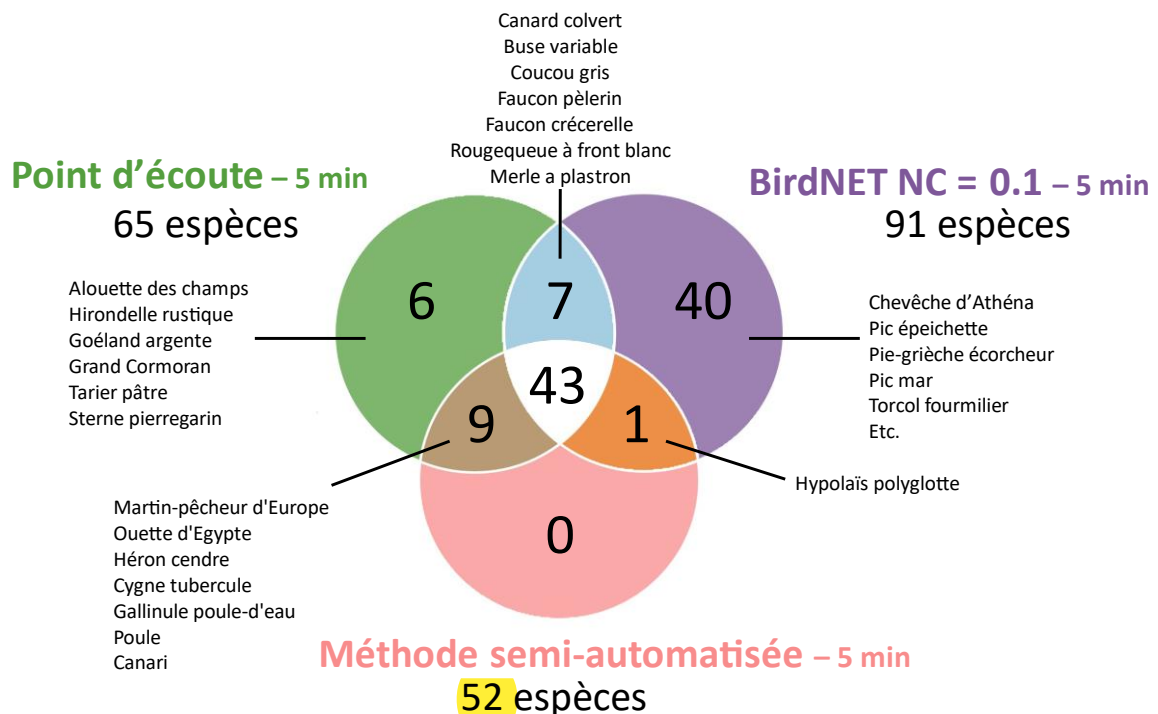


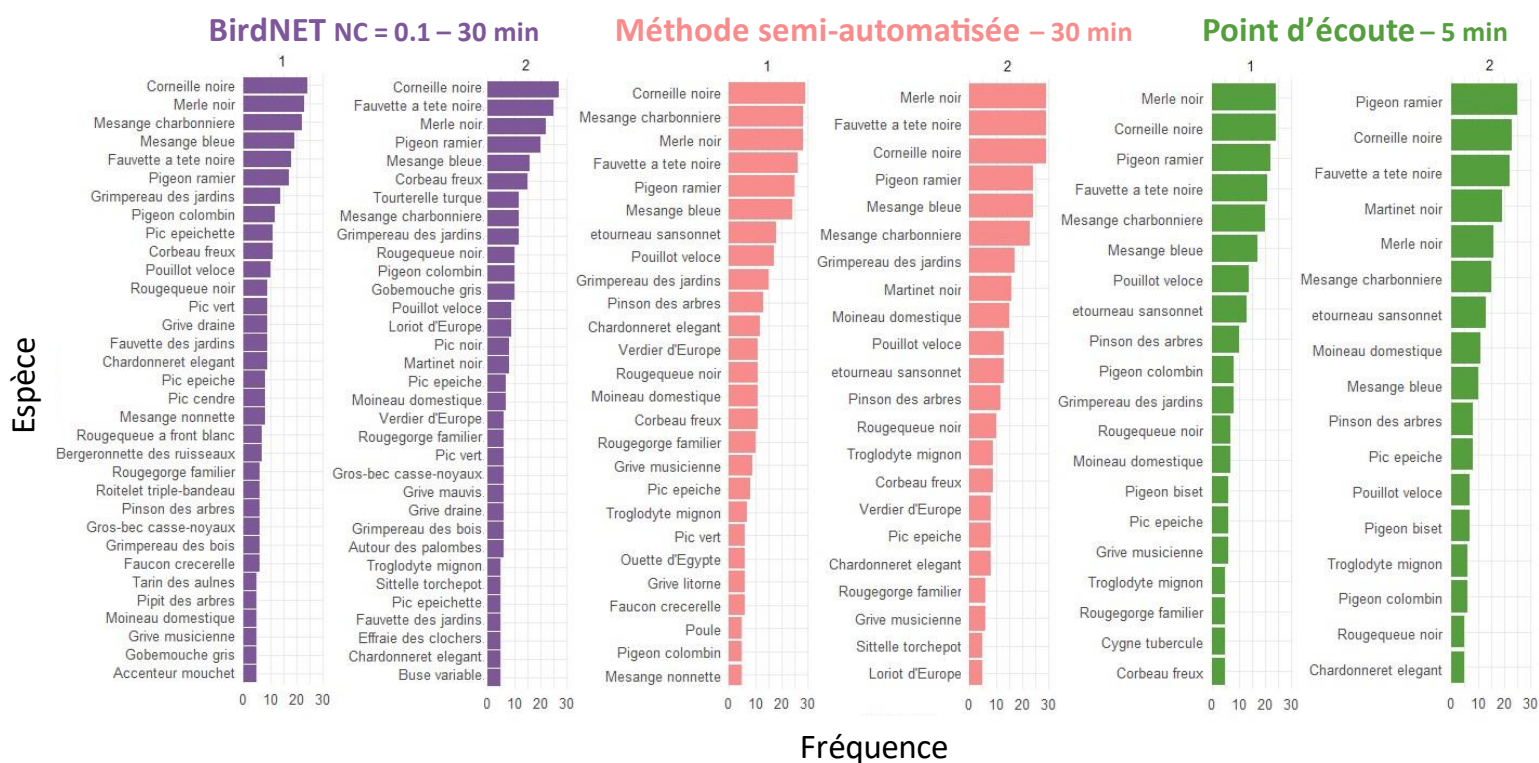
Figure 4 : Diagramme de Venn comparant les communautés obtenues à l'analyse des mêmes 5 min d'enregistrement, sans filtre sur les données de BirdNET (niveau de confiance égal à 0.1).

La méthode semi-automatisée permet de détecter une nouvelle espèce par rapport au point d'écoute : l'hypolaïs polyglotte. Cette méthode permet donc peu d'apport par rapport au point d'écoute sur la même

durée. Enfin, sans filtre sur les données, BirdNET produit quasi autant de FP que TP (40 contre 43), sachant que le statut des 7 espèces communes au point d'écoute et à BirdNET reste flou. En effet, les considérer comme des TP signifierait que BirdNET a été plus précis que l'expert pour la reconnaissance de ces espèces à l'analyse de l'enregistrement (figure 4). Les considérer comme une erreur de BirdNET par rapport à l'enregistrement signifierait une coïncidence curieuse comme ces espèces ont été observées sur les mêmes sites par le point d'écoute. Néanmoins, le grand nombre de FP de BirdNET appelle à la prudence.

1.b. Comparaison des fréquences d'espèces selon les méthodes

Comme l'intérêt d'employer la méthode semi-automatisée et automatisée réside dans l'analyse de durées plus longues qu'un point d'écoute, on s'intéresse à 30 minutes d'enregistrement dans cette partie. Afin d'avoir un aperçu plus détaillé des communautés décrites selon les différentes méthodes, il est possible de comparer les fréquences des espèces. L'histogramme nous indique que les espèces les plus fréquentes sont retrouvées au travers des méthodes, passages et durées : Merle noir (*Turdus merula*), corneille noire (*Corvus corone*), fauvette à tête noire (*Sylvia atricapilla*), mésange charbonnière (*Parus major*) et pigeon



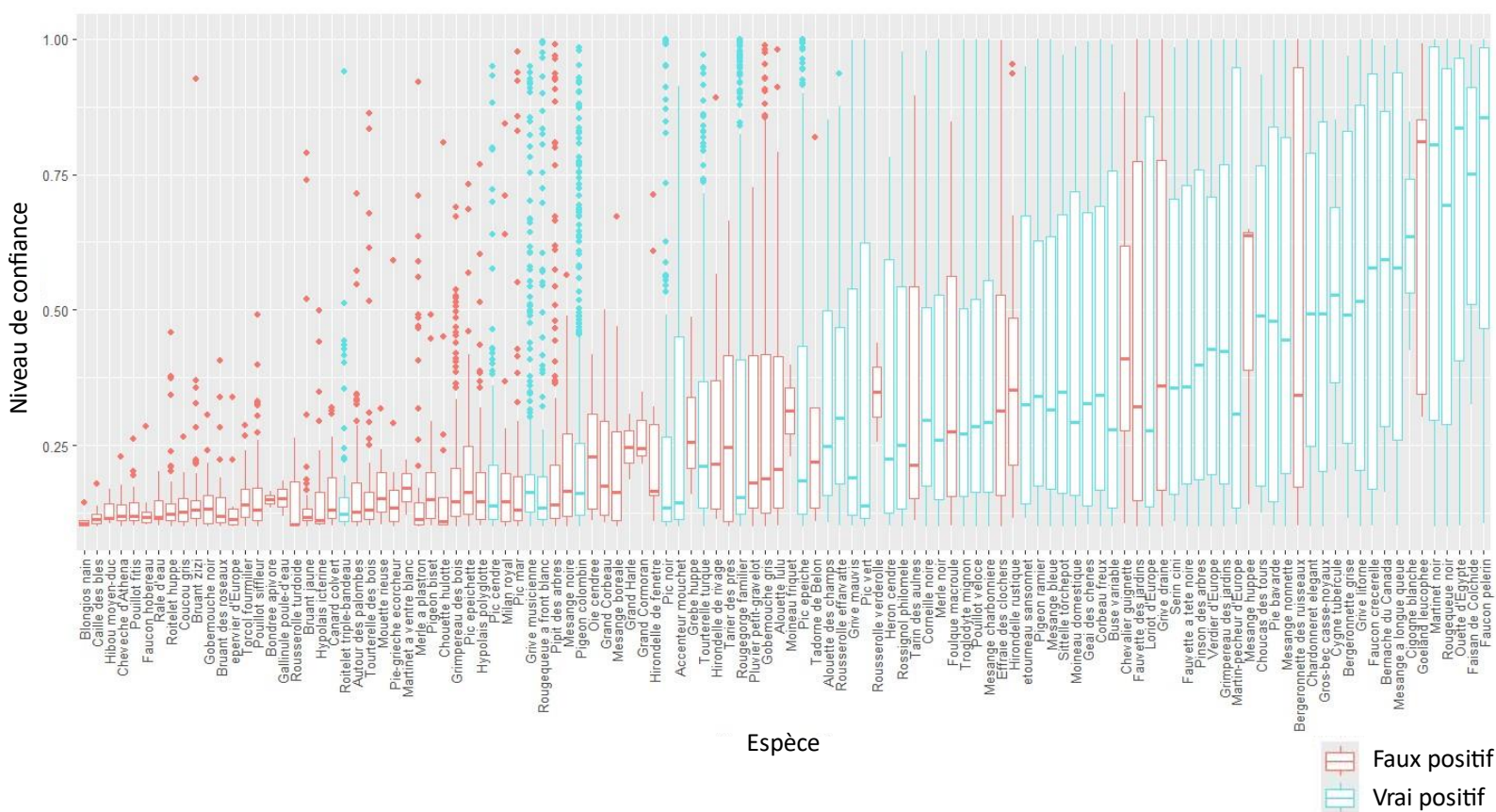
ramier (*Columba palumbus*) sont présents sur au moins 10 sites, ce qui est en cohérence avec leur forte présence en ville.

Figure 5 : Fréquences des espèces sur l'ensemble du relevé pour chaque méthode et passage. Le graphe n'intègre que les espèces présentes sur au moins 5 sites. Le point d'écoute concerne une durée de 5 min par site, les données sont issues d'un enregistrement de 30 min pour la méthode semi-automatisée et automatisée, avec pour cette dernière un niveau de confiance minimal (NC = 0.1).

On peut noter les écarts de diversité alpha pour les différentes méthodes, liés à l'écart de durée d'une part, et à la non sélectivité de la méthode automatisée pour un niveau de confiance fixé à 0.1, comme des espèces sont détectées à tort (Effraie des clochers, Grive mauvis et grive draine) par BirdNET (figure 5). Ainsi, on obtient un premier indice quant à l'importance de la durée d'enregistrement et du filtre de données pour la représentativité de la communauté d'espèces.

1.c. L'hétérogène répartition des niveaux de confiance en fonction des espèces

L'établissement d'un filtre de données via le choix d'un niveau de confiance seuil est un enjeu de l'étude. Ce niveau de confiance unique pour les communautés questionne la variabilité spécifique, (figure 6). Ce



graphe permet également de comparer la disparité des niveaux de confiance associés aux FP aux niveaux de confiance associées aux espèces effectivement présentes (TP) dans le relevé de référence sur les sites où l'espèce a été détectée (méthode semi-automatisée). Bien que les TP bénéficient de niveaux de confiance plus élevés, certains FP sont aussi associés à des niveaux de certitudes forts, dont on peut penser que les chants sont peu distinguables entre espèces appartenant au même genre (goéland leucophaée, bergeronnette des ruisseaux (*Motacilla cinerea*), mésange huppée (*Lophophanes cristatus*)) et effectivement présentes sur les sites. Certains FP dont les niveaux de confiance sont élevés sont effectivement très peu vraisemblables (Chevalier Guignette (*Actitis hypoleucos*), Effraie des clochers (*Tyto alba*), Tadorne de Belon (*Tadorna tadorna*)).

Figure 6 : Niveaux de confiance associées aux points de contacts de chaque espèce du relevé de BirdNET pour 30 min d'enregistrement et un niveau de confiance minimal (0.1). Les espèces sont ordonnées selon la moyenne des niveaux de confiance associées à leur détection.

Ce constat renvoie au choix de la liste des espèces fournie à BirdNET, entre lesquelles l'algorithme choisit les possibles identifications, dont il faudrait adapter le contenu pour s'assurer que ces espèces invraisemblables sont exclues, puisque le filtrage par le niveau de confiance ne suffira pas à les omettre. Enfin, quelques TP ont des niveaux de confiance moyens faibles (Roitelet triple bandeau (*Regulus ignicapilla*), Pic cendré (*Picus canus*), grive musicienne (*Turdus philomelos*), pigeon colombin (*Columba oenas*)), bien que certaines de leurs détections aient des niveaux de confiance très élevés (supérieurs à 0.8) (figure 6). Ceci indique des espèces plus difficiles à détecter et éventuellement faciles à perdre des relevés comme leur détection tient en quelques point de contact avec une certitude forte.

Cela souligne aussi l'importance d'échantillonner une durée suffisante et des sonogrammes assez clairs pour une identification à fort niveau de confiance.

2. Conséquences de l'allongement de la durée d'enregistrement sur la performance

L'augmentation de la durée des enregistrements fournis à BirdNET a deux effets : sur le niveau de confiance i.e. filtre à utiliser pour obtenir une performance optimale, et sur la performance en elle-même, décrite par des indicateurs de performance. Dans l'objectif de bénéficier de cet avantage majeur de la méthode automatisée, les analyses de BirdNET à un niveau de confiance minimal (0.1) pour différentes durées d'enregistrement ont été comparées à celles de l'expert. Ceci sert de base de réflexion pour l'utilisation de matinées d'enregistrement.

2.a. Effet de la durée d'enregistrement sur les performances de BirdNET

Les courbes PR comparent les modèles de classifieur pour chaque durée en mesurant l'évolution du compromis entre *precision* et *recall*. L'AUC augmente avec la durée d'enregistrement ce qui se traduit sur la figure 7 par l'empilement des courbes vers le coin haut-droit et se confirme sur la figure 8. Cette tendance indique une amélioration de la performance avec l'augmentation de la durée d'enregistrement. Le gain de performance est marqué de 5 min jusqu'à 20 min d'enregistrement, au-delà desquels l'AUC atteint un plateau (figure 8). Ceci s'explique par le fait que l'accumulation des enregistrements de 5 min s'est faite par chronologie croissante sur les 2 matinées d'écoute. Ainsi, l'allongement de la durée d'enregistrement permet de détecter l'ensemble des espèces d'oiseaux à l'échelle d'une matinée alors qu'elles commencent à chanter de manière échelonnée. Ainsi l'écart de performance se réduit à partir de 20 min comme le nombre de nouvelles espèces correctement détectées (TP) se réduit. D'autre part, il est raisonnable de penser que le nombre d'espèces nouvellement détectées se réduit également.

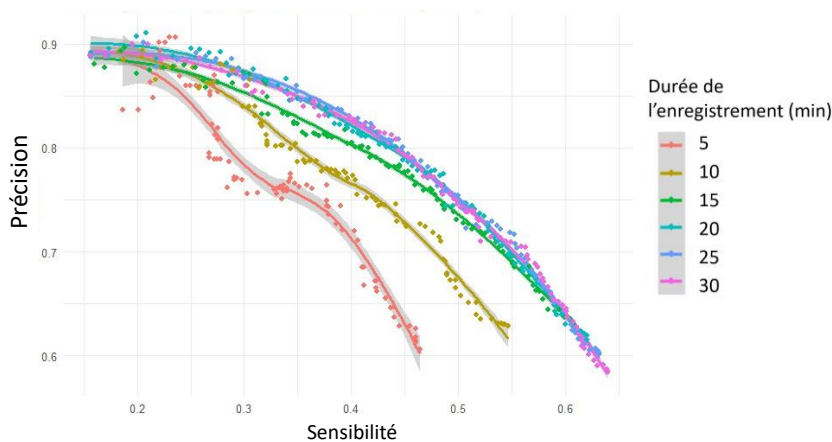


Figure 7 : Courbe PR figurant la précision en fonction de la sensibilité pour différentes durées d'enregistrement. Chaque point représente la moyenne de la précision et du recall de l'ensemble des sites considérés pour un niveau de confiance compris entre 0.10 et 1.

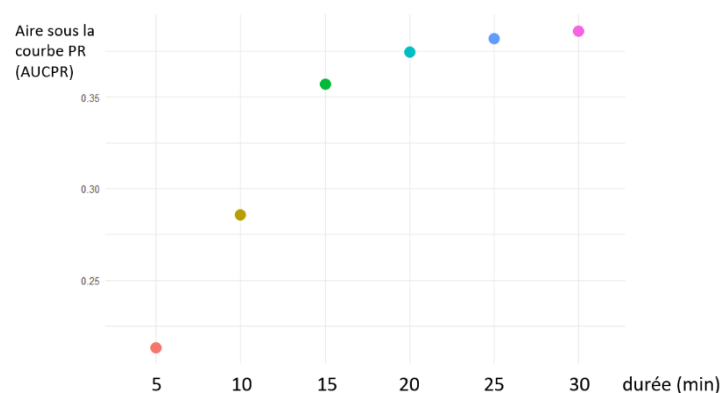


Figure 8 : Aire sous les courbes PR en fonction de la durée de l'enregistrement.

Ainsi, augmenter le temps d'enregistrement contribue à améliorer la performance de l'algorithme, pour un filtrage des détections de BirdNET plus sévère.

2.b. Effet de la durée d'enregistrement sur le niveau de confiance (i.e. filtre de détections de BirdNET)

Les niveaux de confiance ne renseignent pas directement sur la performance de BirdNET. Il est néanmoins intéressant d'évaluer l'évolution des fréquences d'observations pour chaque niveau de

confiance, comme cela peut amener de la variabilité dans le filtre utilisé en fonction de la durée analysée. Une première observation concernant la répartition des niveaux de confiance des détections est l'hétérogénéité de BirdNET. En effet, l'algorithme produit environs 4 fois plus de détections avec des niveaux de confiance très faibles (0.1-0.15 ; 0.15-0.20) que des détections avec des niveaux de confiance moyens à plus élevés (0.35-0.9) quelle que soit la durée d'enregistrement (figure 9 gauche). Ceci montre la propension de BirdNET à la production d'une masse de FP nuisant à la véracité de l'analyse. Cependant, on observe aussi une grande quantité de détections avec une certitude élevée (niveau de confiance compris entre 0.9 et 1) (figure 9 gauche). Le choix d'un filtre revient donc à éliminer une masse d'informations avec des niveaux de certitude faibles, mais la question de la valeur de ce filtre se pose.

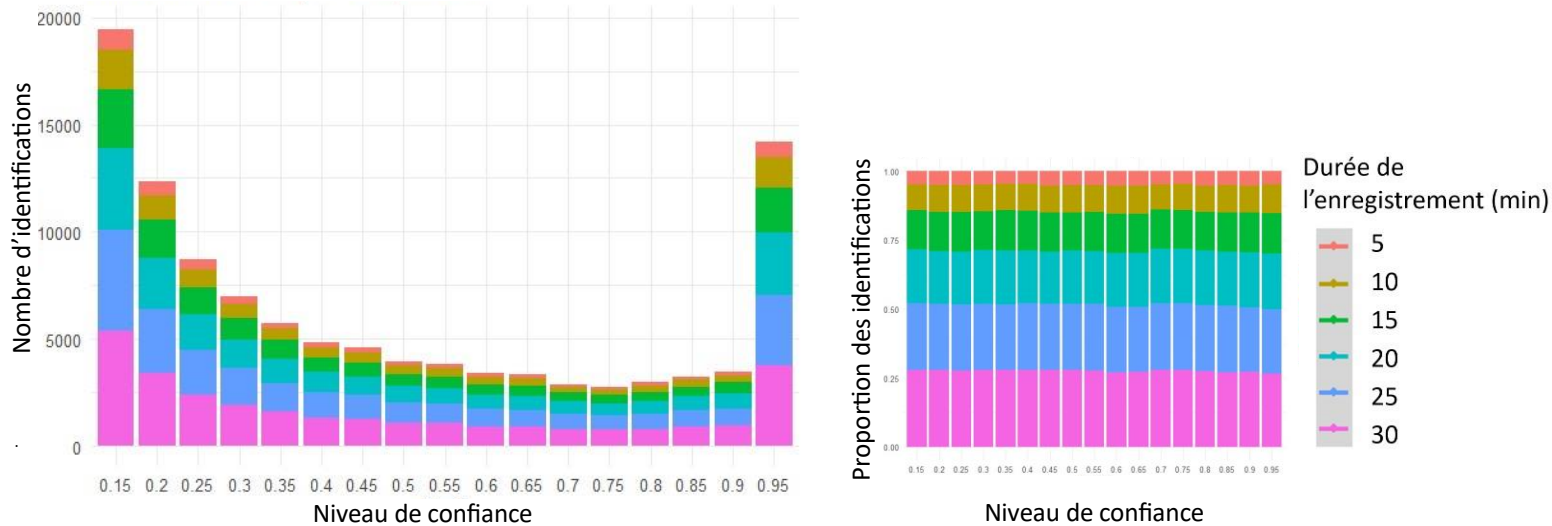


Figure 9 : Fréquence des catégories de niveaux de confiance associées aux détections pour différentes durées d'enregistrement (gauche) et proportion d'identifications pour les différentes durées pour chaque catégorie de niveau de confiance (droite). Les catégories correspondent à des bins valant 0.05 unités de niveau de confiance.

Ainsi on voit apparaître le compromis entre la sauvegarde des informations correctes et des erreurs. Elever la durée d'enregistrement augmente le nombre de contacts pour chaque classe de niveaux de confiance (figure 9 droite). On peut donc espérer établir un filtre de valeur plus importante (donc plus sévère) avec l'augmentation de la durée, en parallèle de l'amélioration de la performance de l'algorithme BirdNET.

3. Le filtrage des données

3.a. Différents filtres correspondant à différents aspects de la performance

Pour le choix d'un filtre, i.e. une valeur de niveaux de confiance minimum pour retenir une identification d'espèce, l'association des différentes métriques de quantification de la performance de BirdNET est essentielle pour expliquer les sources d'évolution de la performance entre niveaux de confiance. Dans le but d'appliquer la méthode à des matinées entières d'enregistrement, cette section compare la performance de BirdNET à 30 min à celle de l'analyse de l'expert sur les mêmes enregistrements.

La *precision* ou taux de détections correctes croît au détriment de la sensibilité (*recall*) pour des niveaux de confiance plus élevés (Figure 10). L'indice de similarité de Jaccard des communautés détectées par BirdNET et l'expert à l'écoute de l'enregistrement est maximal pour un filtre à NC = 0.38, valant alors environ 0.45. La similarité pouvant être écrite $TP / (TP + FP + FN)$, cette situation renvoie à une sensibilité où environ 55% des espèces qui devraient être détectées le sont, tandis que 70 % des détections sont correctes. Ce filtre correspond aussi à celui pour lequel le F1score est maximal, i.e. la moyenne harmonique équilibrant la *precision* et le *recall*, ainsi que le F0.75score, donnant plus de poids

à la *precision*. La courbe PR renvoie également à ce niveau de confiance-ci, optimisant la distance au point (1,1), point pour lequel la performance de l'algorithme est maximale tout en équilibrant *precision* et *recall* (figure 11).

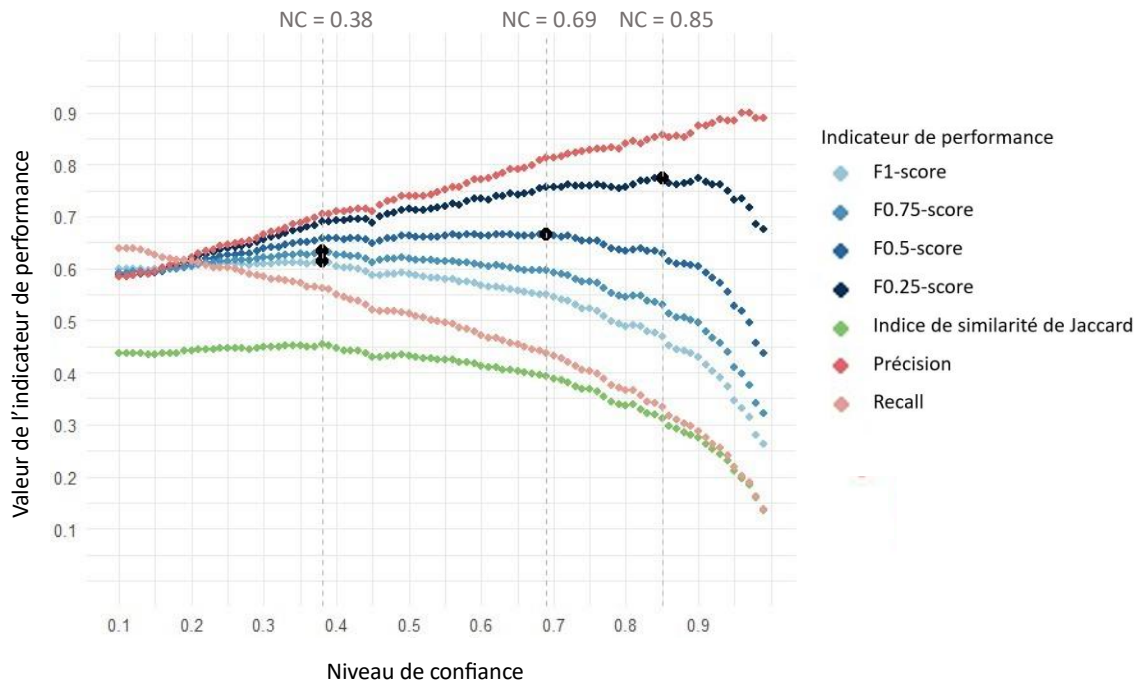


Figure 10 : Valeur des indicateurs de performance en fonction du niveau de confiance sélectionné. Chaque point correspond à la moyenne pour l'ensemble des sites considérés sur 30 min d'enregistrement.

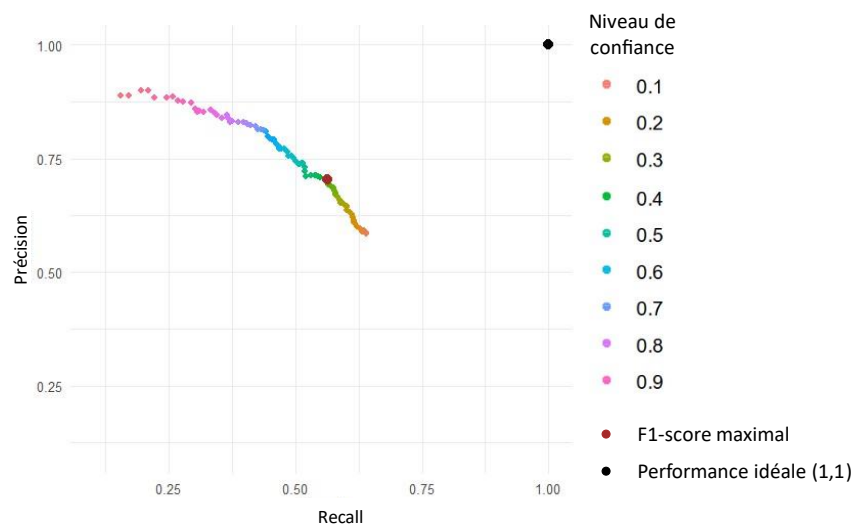


Figure 11 : Courbe PR avec le point correspondant au maximum de la moyenne harmonique de précision et recall. Chaque point correspond à la moyenne pour l'ensemble des sites considérés sur 30 min d'enregistrement.

D'autres filtres sont également préconisés par des indicateurs différents. Le F0.5-score et F0.25-score, donnant un poids accru à la *precision*, atteignent leur maximum respectivement pour NC = 0.69 et NC = 0.85. Ainsi, le pourcentage de détections correctes (*precision*) vaut respectivement à 0.8, puis 0.85 au détriment de la sensibilité qui chute alors à 0.45 et 0.35 d'espèces effectivement détectées.

3.b. Comparaison de ces filtres et des communautés qu'ils produisent

En se plaçant à l'échelle de l'ensemble des sites considérés, on peut suivre l'évolution du tableau de contingence global pour ces trois niveaux de confiance. Les TP regroupent les espèces correctement détectées sur les sites, les FP les espèces toujours détectées à tort sur les sites et celles qui devraient être totalement absentes du relevé et les FN les espèces qui ne sont jamais détectées sur les sites où elles devraient l'être, ou totalement absentes des détections de BirdNET à tort. Le nombre global de TP et FN reste stable de NC = 0.38 à NC = 0.69, tandis que le nombre de FP chute drastiquement, passant de 42 à 26 espèces détectées à tort (figure 12). Comme ce graphe se place à une échelle globale, il ne montre pas le nombre de TP perdus par site et donc le nombre de FN augmentant à l'échelle des sites, tant qu'il y a au moins un site où l'espèce concernée est correctement détectée. L'intérêt de ce graphe réside néanmoins dans la vue globale qu'il permet d'avoir sur l'ensemble des sites avec le nombre d'espèces total correctement détectées et ce qui est perdu avec des filtres plus sévères. Ce graphe appuie l'importance de considérer des niveaux de confiance plus élevés comme ceux-ci contribuent à réduire drastiquement le nombre de FP, sans faire exploser le nombre de FN.

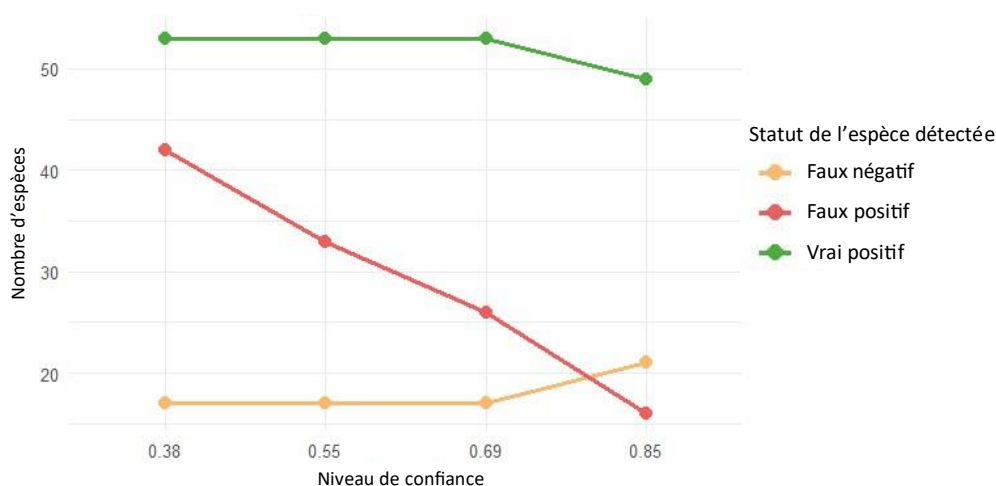


Figure 12 : Evolution du tableau de contingence pour 30 min d'enregistrement analysées en fonction du niveau de confiance seuil. Les valeurs correspondent aux totaux pour l'ensemble des sites. Les valeurs pour NC = 0.55 sont explicatives et permettent d'étoffer le graphe.

Quel que soit le niveau de confiance utilisé pour filtrer les données, les communautés obtenues avec BirdNET par rapport à celles issues de l'analyse de l'enregistrement par l'expert sont différentes (similarité moyenne ne dépassant pas 45%), ce qui est confirmé par une ANOSIM. La NMDS nous montre cependant que les différences entre les communautés obtenues via BirdNET s'accroissent avec un niveau de confiance croissant (figure 13), résultant en un écart plus grand entre les sites qui s'observe dans la valeur R positive et croissante, marquant une différence inter-méthode en augmentation. On observe que les communautés des sites ont aussi tendance à se distinguer à mesure que l'on augmente le niveau de confiance, se traduisant par l'agrandissement de l'ellipse bleue correspondant aux détections de BirdNET par rapport à l'ellipse orange liée aux communautés détectées par l'expert dans l'enregistrement. Néanmoins, les centroïdes demeurent proches, indiquant une cohérence entre les communautés décrites par les 2 méthodes.

Ainsi, bien que les différences s'accroissent (le nombre de FN par site augmente) entre la méthode semi-automatisée et automatisée avec l'augmentation du niveau de confiance, la chute du nombre de FP contribue à une amélioration de la précision de l'algorithme.

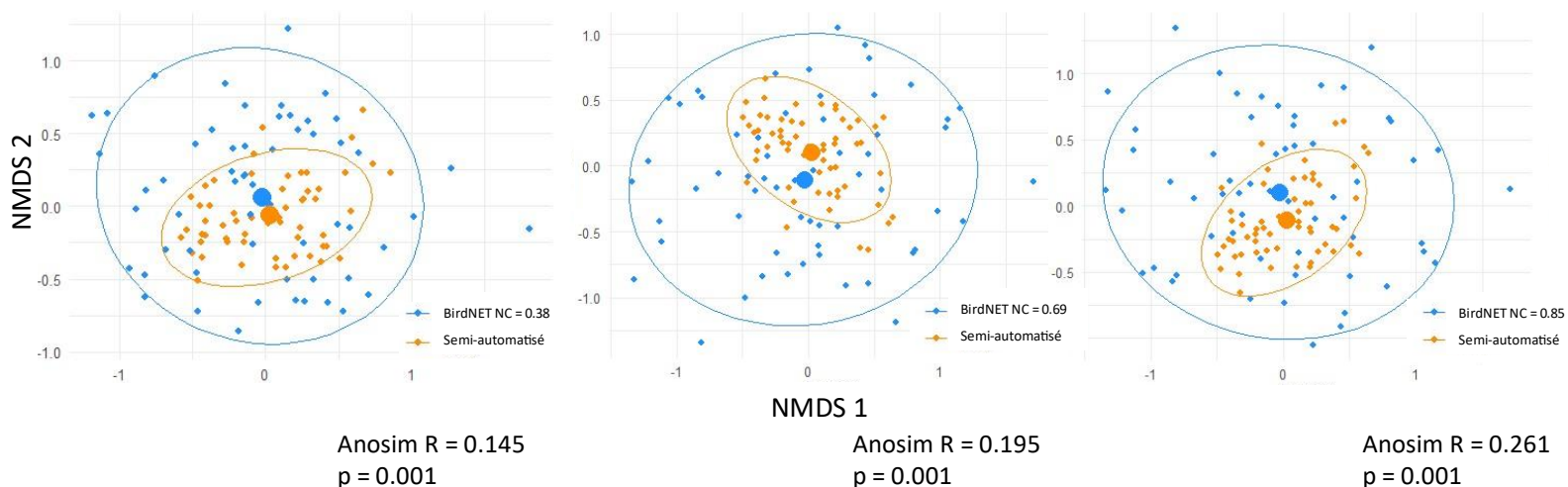


Figure 13 : Comparaison par NMDS des communautés obtenues avec l'enregistrement et BirdNET pour différentes valeurs de niveau de confiance. Chaque point figure un site, les 2 points plus gros sur chaque graphe représentent les centroïdes. Les ellipses et les centroïdes figurent les NMDS et valeur des ANOSIM permettent ici de comparer les différences inter-communautés. Les enregistrements de 30 min sont considérés.

4. Vers des matinées d'enregistrement : le choix d'un filtre

4.a. Validité de la méthode & des filtres pour des durées plus longues

L'analyse de la performance pour 6x5 min d'enregistrement réparties sur 2 matinées d'écoute a permis de dégager trois niveaux de confiance correspondant à des maxima de performance. Etudier ces 3 maxima et les filtres associés pour différentes durées d'enregistrement permet d'évaluer dans quelle mesure il est possible de les projeter sur des enregistrements plus longs, de « matinées » (de 15 min avant le lever du soleil à 9h du matin environ, soit 3 à 4h en fonction de si l'on est en avril ou en juin)

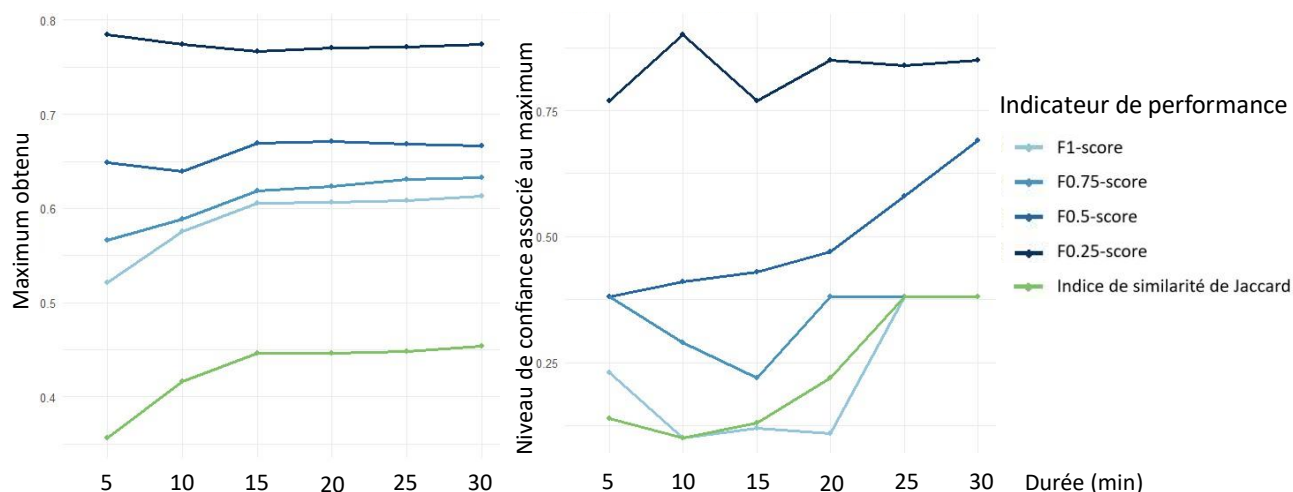


Figure 14 : Evolution des valeurs moyennes des maxima d'indicateurs de performance (gauche) et des valeurs moyennes de niveaux de confiance associées (droite) en fonction de la durée des enregistrements. Chaque point correspond à une moyenne pour l'ensemble des sites considérés.

Dès 3x5 min d'enregistrement un plateau est atteint pour les maxima de performance (figure 14). Il est alors attendu de retrouver ces valeurs pour des durées plus longues d'enregistrement.

Pour ce qui est des filtres associés à ces maxima, les valeurs sont plus variables. Pour le F0.25-score et le F0.75-score, une stabilisation semble se dessiner à partir de 20 min d'enregistrement, avec des valeurs de filtre respectivement à 0.85 et 0.39. Pour le maximum de l'indice de similarité de Jaccard et le F1-score, la valeur du niveau de confiance associé est inchangée entre 25 et 30 min d'enregistrement. Enfin, le niveau de confiance associé au F0.5-score croît à chaque ajout de 5 min d'enregistrement et on peut s'attendre à ce que sa valeur soit différente pour des durées longues d'enregistrement. Ainsi, sur la durée d'une matinée, on peut s'attendre à retrouver des valeurs équivalentes pour les maxima des niveaux de confiance, mais les valeurs de niveau de confiance associées seront probablement différentes et certainement plus élevées. Néanmoins, cette analyse ne permet pas de dégager des filtres optimaux sur des durées d'enregistrement plus longues, même si elle pousse à s'intéresser aux communautés décrites pour des filtres « sévères ».

4.b. Comparaison de filtres pour des durées longues d'enregistrement

Ainsi, on applique une méthode conservative en utilisant les filtres NC = 0.69 et 0.85 aux enregistrements des « matinées ». On applique également les filtres suivants : 0.90, 0.95, afin de parer à l'augmentation du nombre de FP produite par l'accumulation de données et donc de maintenir une précision élevée. La seule façon d'avoir un aperçu de la performance pour ces deux derniers niveaux de confiance consiste à comparer les analyses de l'algorithme avec celles produites à NC = 0.69 et NC = 0.85. On ne reprend pas le filtre NC = 0.38 comme l'accumulation de données conduit à un nombre aberrant de FP pour ce filtre.



Figure 15 : Evolution de la taille du relevé de BirdNET en fonction de différents niveaux de confiance sur les sites de l'étude pour les enregistrements de matinées. Les espèces perdues sont mentionnées.

En accentuant la sévérité du filtre, le nombre d'espèces détectées à l'échelle de l'ensemble du relevé se réduit passant de 110 à 69 (figure 15). Ne possédant pas la liste de référence des espèces détectées pour des durées d'enregistrement si longues, il n'est pas possible de connaître toutes les erreurs et succès de l'analyse de BirdNET. En revanche, la liste des espèces entendues à l'écoute de 30 min d'enregistrement

et sur les points d'écoute peut servir de pseudo-référence, non exhaustive. Ainsi, les listes d'espèces de la figure ... représentent les espèces éliminées du relevé et parmi celles-ci les espèces en orange font partie de la liste de pseudo-référence et sont perdues alors que leur présence sur les sites est certaine. Certaines de ces espèces perdues sont rarement observées sur les sites, comme l'alouette des champs (*Alauda arvensis*), le pic cendré, la linotte mélodieuse (*Linaria cannabina*), l'hypolaïs polyglotte (*Hypolaïs polyglotta*), le tarier des prés (*Saxicola rubetra*), l'hirondelle de fenêtre (*Delichon urbicum*) ou la rousserolle verderolle (*Acrocephalus palustris*) ce qui favorise leur élimination du relevé. D'autres, telles que la cigogne blanche (*Ciconia ciconia*), la bernache du Canada (*Branta canadensis*) ou la gallinule poule d'eau (*Gallinula chloropus*) émettent plutôt des cris que des chants et sont ainsi assez silencieuses, rendant difficile leur détection par l'algorithme. Cette figure montre de manière globale que les espèces d'eau, ne chantant pas, sont aisément éliminées du relevé entre $NC = 0.69$ et $NC = 0.85$. Ceci peut être problématique mais nombre d'entre elles sont improbables sur les sites d'étude (Bihoreau gris (*Nycticorax nycticorax*), blongios nain (*Ixobrychus minutus*), canard chipeau (*Mareca strepera*), râle d'eau (*Rallus aquaticus*), pluvier petit-gravelot (*Charadrius dubius*)). On note également la perte du Goéland leucopée.

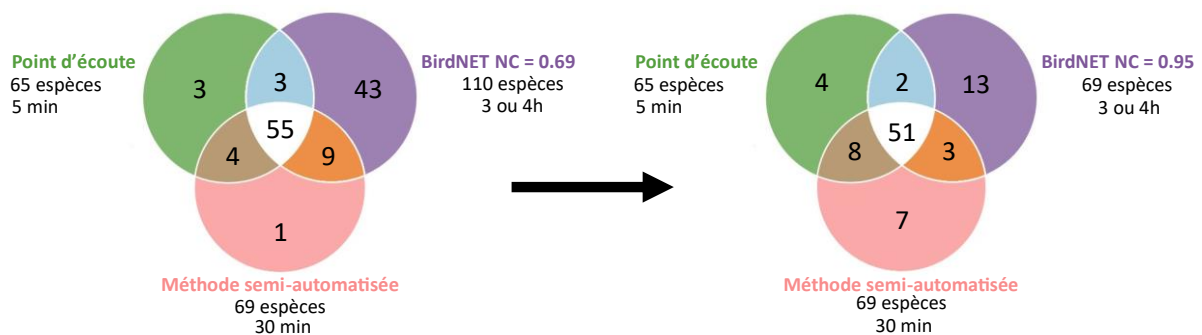


Figure 16 : Comparaison des diagrammes de Venn pour $NC = 0.69$ et $NC = 0.95$ résumant le nombre d'espèces détectées selon les méthodes. Le temps d'enregistrement ou d'écoute est précisé sur la figure.

Les diagrammes de Venn (figure 16) appuient l'importance d'appliquer un filtre sévère pour des durées longues. De la même manière que précédemment, ces diagrammes sont indicatifs comme l'on ne possède pas une liste de référence intégrale des espèces observées sur une matinée. Ainsi il serait incorrect de considérer l'ensemble des espèces obtenues uniquement par BirdNET comme des FP. Néanmoins, il est peu probable que le nombre d'espèces détectées par BirdNET excède grandement celui obtenu via les autres méthodes et durées, d'autant plus que BirdNET produit des faux négatifs. D'après ces diagrammes, il semble intéressant d'utiliser un niveau de confiance à 0.9 ou 0.95.

4.c. Vers le choix d'un filtre

Pour trancher entre ces deux niveaux de confiance, on peut s'appuyer sur la NMDS comparant les communautés des sites pour $NC = 0.9$ et 0.95 . Les centroïdes des communautés décrites par méthode sont quasiment confondus, indiquant une similarité forte entre les méthodes. D'autre part, l'ANOSIM ne permet pas de conclure à la présence d'une différence significative entre les communautés selon les méthodes pour $NC = 0.9$ et 0.95 , traduite sur la NMDS (figure 17) par la proximité des ellipses pour chaque méthode. Pour des résultats similaires entre deux niveaux de confiance, il convient donc de prendre le plus élevé pour améliorer la précision de la méthode, soit $NC = 0.95$.

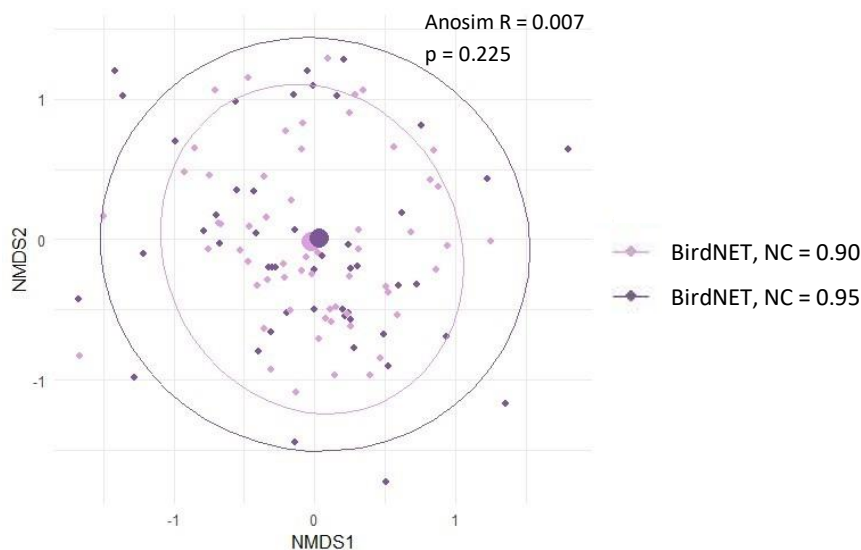


Figure 17 : Comparaison des communautés obtenues sur les sites grâce à BirdNET sur des enregistrements de matinées en employant deux niveaux de confiance : 0.90 et 0.95. Les points correspondent aux sites et les points plus gros sont les centroïdes des ellipses. On considère les enregistrements des « matinées ».

5. Ecologie

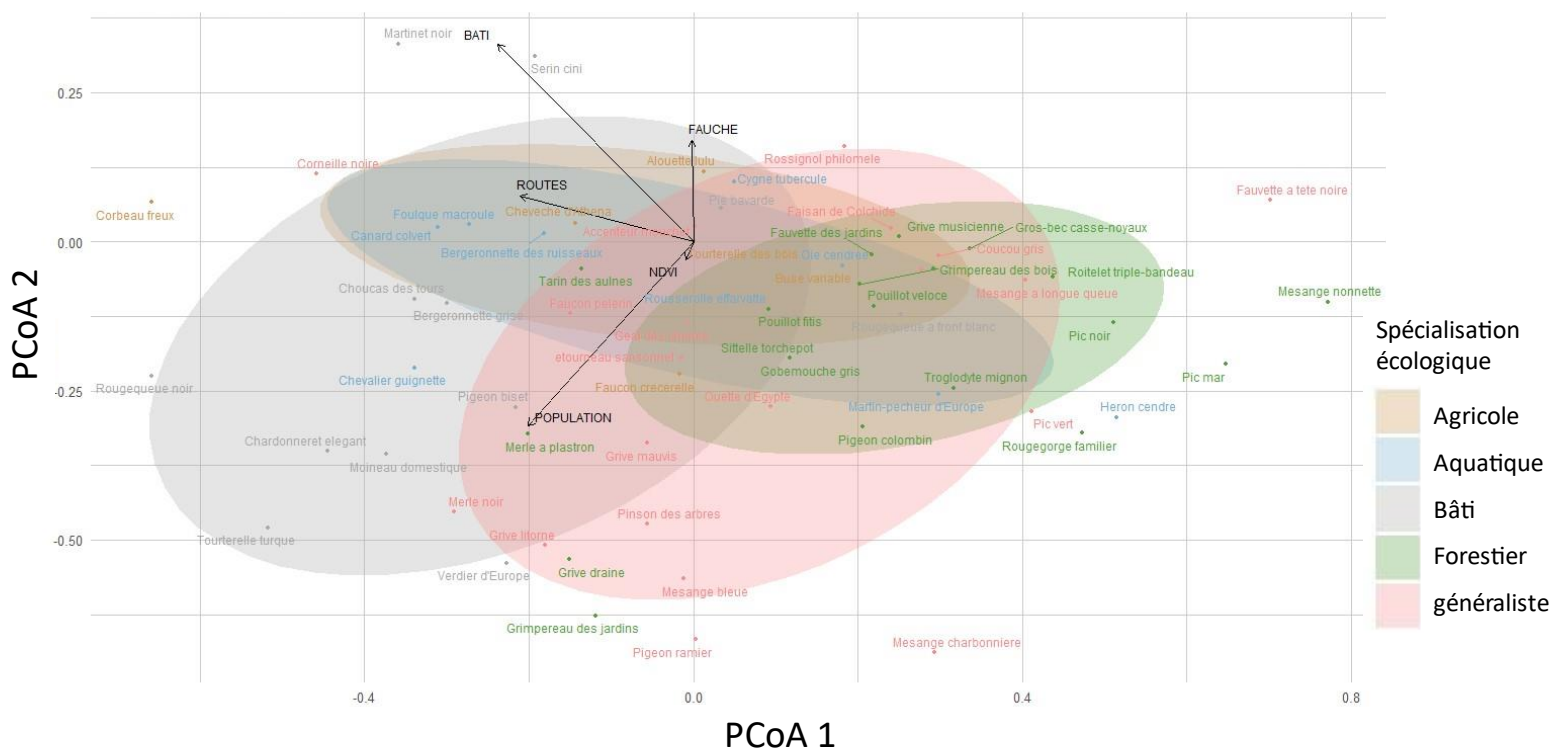


Figure 18 : Répartition des espèces détectées par BirdNET sur les enregistrements de matinées et des caractéristiques paysagères des sites selon les deux premiers axes de la PCoA. Les espèces ont été colorées selon leur spécialisation écologique et les ellipses regroupent selon ces mêmes spécialisations 70% des espèces.

Les ellipses regroupent 70 % des espèces d'une spécialisation écologique (figure 18). On observe que les spécialistes des habitats forestiers se regroupent et sont anticorrélées avec des variables liées aux milieux urbains comme la compacité du bâti ou la densité de routes. Les spécialistes des milieux urbains

sont corrélées à ces variables-là et à la variable de densité de population, à l'opposé des variables forestières et leur ellipse se distingue alors relativement bien de l'ellipses des spécialistes des milieux forestiers. Malgré la cohérence de ces résultats, on observe que la variable de végétalisation (NDVI, *normalized difference vegetation index*) ne semble pas décrire les communautés forestières. Ces résultats très qualitatifs ne permettent pas d'affirmer que les données de BirdNET de cette étude seront assez précises et correctes pour établir des résultats écologiques pertinents, mais ils poussent à être optimiste quant à la méthode.

6. L'étude des biais de la performance de l'algorithme

6.a. Caractérisation de la pollution sonore sur les sites

Le niveau d'intensité sonore moyen sur les sites de l'étude est groupé surtout entre 30 et 40 dB, soit un environnement sonore calme (figure 19). Quelques sites plus bruyants pour lesquels le niveau d'intensité sonore monte jusqu'à 51dB sont à noter.

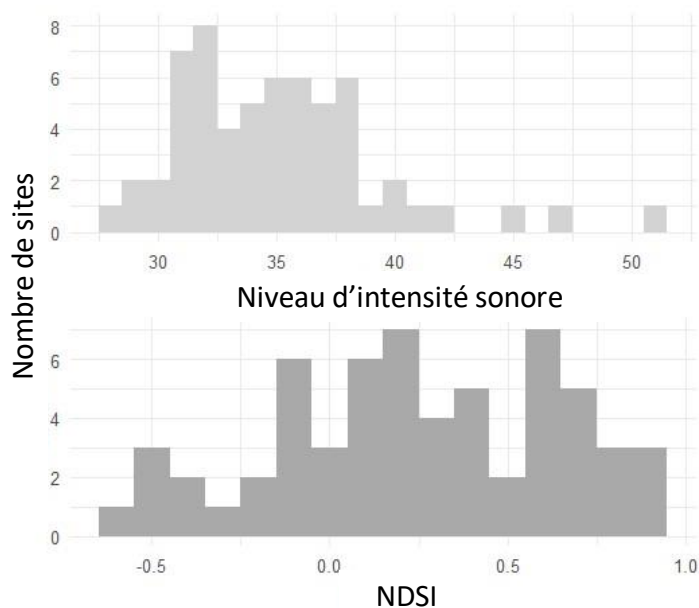


Figure 19 : Répartition des sites du relevé selon le niveau d'intensité sonore moyen et du NDSI. Les données sont des moyennes par site pour chaque passage sur 30 minutes d'enregistrement.

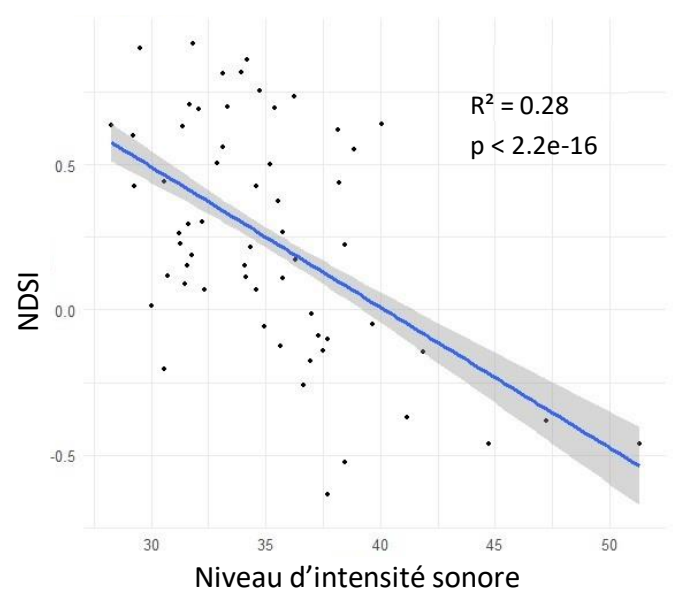


Figure 20 : Lien entre le niveau d'intensité sonore moyen et NDSI. Chaque point correspond à la moyenne des valeurs obtenues pour 30 min d'enregistrement sur chaque site pour les 2 passages.

Pour ce qui est du NDSI, La majorité des sites ont une dominante biophonique dans le paysage sonore, qui peut être très marquée avec des valeurs de NDSI s'approchant de 1, alors que quelques sites ont une composante anthropophonique dominante pour des valeurs négatives de NDSI. Le NDSI est négativement corrélé au niveau d'intensité sonore (figure 20), indiquant que les sites les plus bruyants sont ceux dont le paysage sonore est plutôt dominé par l'anthropophonie. Cette relation est validée par un R^2 relativement marqué signifiant que la relation explique 28% de la variance.

6.b. Lien entre les variables de pollution sonore et les caractéristiques paysagères des sites

6.b.i. Niveau d'intensité sonore et variables paysagères

La moyenne du niveau d'intensité sonore sur 30 min d'enregistrement augmente fortement avec la densité de routes dans un rayon de 150m (tableau 4 ; figure 21). D'autres variables ont montré un effet

significatif, mais la taille de l'effet estimée est négligeable par rapport à celle de la densité de routes. Ceci permet de pointer le trafic routier en tant que source de pollution sonore majeure sur nos sites.

Variable explicative	Estimation de l'effet	Valeur de la statistique (t value)	Df	P-value
Intercept	3.23e1 ±1.24	26.14	342	< 2e-16
Densité de routes dans un rayon de 250m	2.40e+1 ±2.1e+1	11.404	342	< 2e-16
Fréquence de fauche annuelle	-1.90e-2 ±5.31e-2	-0.358	342	0.72
Densité de population dans un rayon de 150m	-2.22e-4 ±6.33e-5	-3.508	342	5.1e-4
Surface végétalisée dans un rayon de 250m	3.24 ±1.9	1.701	342	8.9e-2
Compacité du bâti dans un rayon de 300m	-8.63e-1 ±3.31e-1	-2.610	342	9.4e-3

Tableau 4 : Description du modèle lm considérant le niveau d'intensité sonore moyen sur 30 min d'enregistrement expliqué par les variables paysagères des sites Evolvie. $R^2 = 0.31$

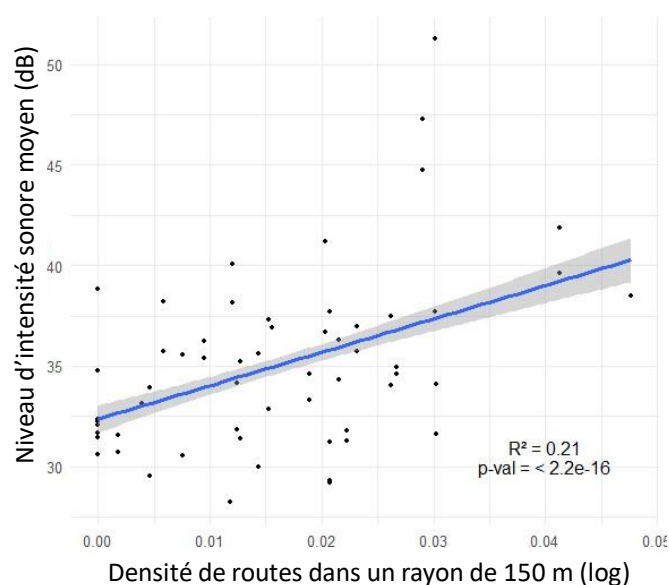


Figure 21 : Lien entre le niveau d'intensité sonore moyen et la densité de routes. Les points représentent la moyenne du niveau d'intensité sonore sur 30 minutes d'enregistrement, en fonction du log de la densité de route dans un rayon de 150 m.

6.b.ii. NDSI et variables paysagères

Pour ce qui est de la nature de la source de son, les valeurs positives de NDSI, i.e. les valeurs marqueurs de biophonie sont liées aux sites dont le recouvrement végétalisé dans un rayon de 250m est élevé (figure 22 ; tableau 5), soit des sites peu urbains. A l'inverse, des valeurs négatives de NDSI (anthropophonie) sont à relier avec des sites pour lesquels la densité de routes est élevée dans un rayon de 150 m (figure 22 ; tableau 5). Le NDSI s'avère être un bon descripteur de du paysage des sites, en cohérence avec sa définition.

Variable explicative	Estimation de l'effet	Valeur de la statistique (t value)	Df	P-value
Intercept	-6.22e-2 ±1.09e-1	-0.57	342	0.57
Densité de routes dans un rayon de 250m	-1.41e+1 ±1.85	-7.65	342	2.14e-13
Fréquence de fauche annuelle	5.25e-3 ±4.67e-3	1.13	342	0.26

Densité de population dans un rayon de 150m	8.25e-6 ±5.56e-6	1.48	342	0.139
Surface végétalisée dans un rayon de 250m	1.14 ±1.67e-1	6.79	342	5.13e-11
Compacité du bâti dans un rayon de 300m	2.25e-2 ±2.90e-2	0.77	342	0.44

Tableau 5 : Description du modèle lm considérant le NDSI moyen sur 30 min d'enregistrement expliqué par les variables paysagères des sites Evolville. $R^2 = 0.39$

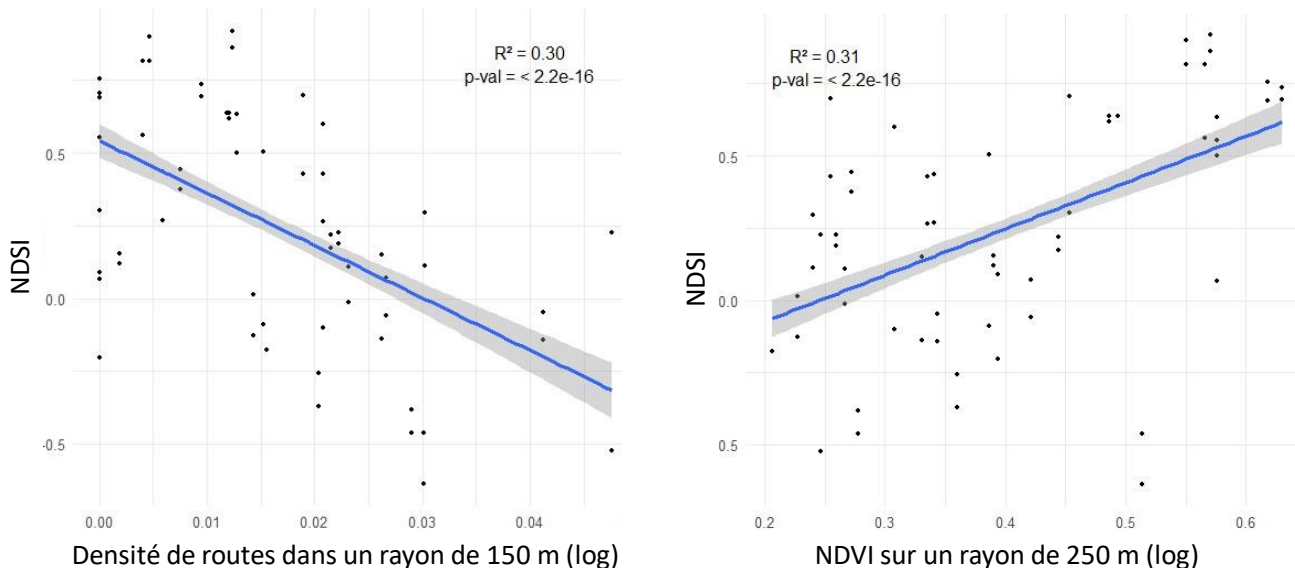


Figure 22 : Relation entre anthropophonie et biophonie et les éléments du paysage sur les sites. Les points représentent la moyenne du NDSI sur 30 minutes d'enregistrement, en fonction du log de la densité de route dans un rayon de 150 m et du log de la surface végétalisée dans un rayon de 250 m.

6.b.iii. Taille des sites et gradient urbain

Par curiosité, on vérifie le lien entre la taille du site et le gradient urbain, que l'on caractérise par deux variables paysagères : la surface végétalisée du site et la densité de routes. Ce lien significatif étoffe la description des sites et leur lien avec le paysage sonore. Ainsi, les sites grands ont un couvert végétalisé dominant. A l'inverse, les sites urbains, à proximité de routes, sont plus petits (tableau 6).

Variable explicative	Estimation de l'effet	Valeur de la statistique (t value)	Df	P-value
Intercept	-1.2e3 ±4.7e3	-0.25	357	0.805
Surface végétalisée dans un rayon de 250m	5.6e+4 ±8.9e3	6.245	357	1.20e-09
Densité de routes dans un rayon de 250m	-5.5e5 ±9.4e-4	-5.845	357	1.14e-8

Tableau 6 : Description du modèle lm étudiant le lien entre la taille du site et 2 variables paysagères : la surface végétalisée et la densité de routes dans un rayon de 250 m. $R^2 = 0.29$

6.c. Lien entre les variables de pollution sonore et performance

Après avoir établi les liens entre l'environnement sonore et le paysage des sites, l'étude s'est intéressée aux niveaux de confiance associés aux détections d'espèces TP, i.e. d'espèces entendue sur le site soit par l'enregistrement soit par le point d'écoute ou les deux, pour étudier le lien entre les variables de

pollution sonore et la performance de l'algorithme. Elle montre que l'effet de la pollution sonore est espèce-spécifique (tableau 7).

Espèce	Pseudo-R ²	variables	Valeur de la statistique (Chisq)	p-value	Df
Grimpereau des jardins	0.50	Imean_6	4.48	0.034	1
		ndsi_6	22.38	2.234e-06	1
Faucon crécerelle	0.76	Imean_6	13.26	0.00027	1
		ndsi_6	2.37	0.12	1
Fauvette a tête noire	0.18	Imean_6	4.67	0.031	1
		ndsi_6	7.95	0.0048	1
Martinet noir	0.57	Imean_6	1.23	0.27	1
		ndsi_6	4.90	0.027	1
Mésange nonnette	0.70	Imean_6	5.14	0.023	1
		ndsi_6	10.26	0.0014	1
Moineau domestique	0.46	Imean_6	9.40	0.0022	1
		ndsi_6	7.56	0.0060	1
Pic épeiche	0.41	Imean_6	4.87	0.028	1
		ndsi_6	4.14	0.041	1
Pigeon ramier	0.21	Imean_6	1.12	0.29	1
		ndsi_6	5.43	0.022	1

Tableau 7 : Description du modèle considérant le niveau de confiance moyen par site, passage et par espèce pour 30 min d'enregistrement en fonction des 2 variables de pollution sonore. De même que précédemment, les espèces sont des vrais positifs sur les sites.

6.c.i. Effet du niveau d'intensité sonore sur le niveau de confiance moyen

Pour le grimpereau des jardins (*Certhia brachydactyla*), le moineau domestique (*Passer domesticus*) et le pic épeiche (*Dendrocopos major*), le niveau d'intensité sonore semble avoir un effet négatif sur le niveau de confiance moyen (figure 23). Cet effet peut s'expliquer par le fait que l'algorithme peine plus à reconnaître le motif de chant sur un spectrogramme saturé, et/ou que l'oiseau a peut-être modifié sa fréquence de chant ou son motif. D'autre part, le pic épeiche émet des cris et non des chants, ce qui peut également contribuer à la difficulté de sa reconnaissance et potentiellement provoquer de la confusion avec d'autres espèces de pic à cause de la similarité de leurs cris. Il est à noter que les nuages de points sont assez dispersés, indiquant que l'effet de cette variable de pollution sonore reste limité voir discutable. De même, le niveau de confiance moyen associé aux détections de la fauvette à tête noire répond de manière peu claire au niveau d'intensité sonore et la relation positive semble tenir aux deux points correspondant aux maxima de niveau d'intensité sonore (figure 23). Cependant, une proposition d'explication pour cette relation positive serait que même dans un environnement sonore bruyant, la fauvette a tendance à dominer (niveau d'intensité sonore du chant élevée) et être facilement reconnue. La faible quantité de données pour le faucon crécerelle et la mésange nonnette pousse à rester prudent sur les conclusions liées à ces deux espèces.

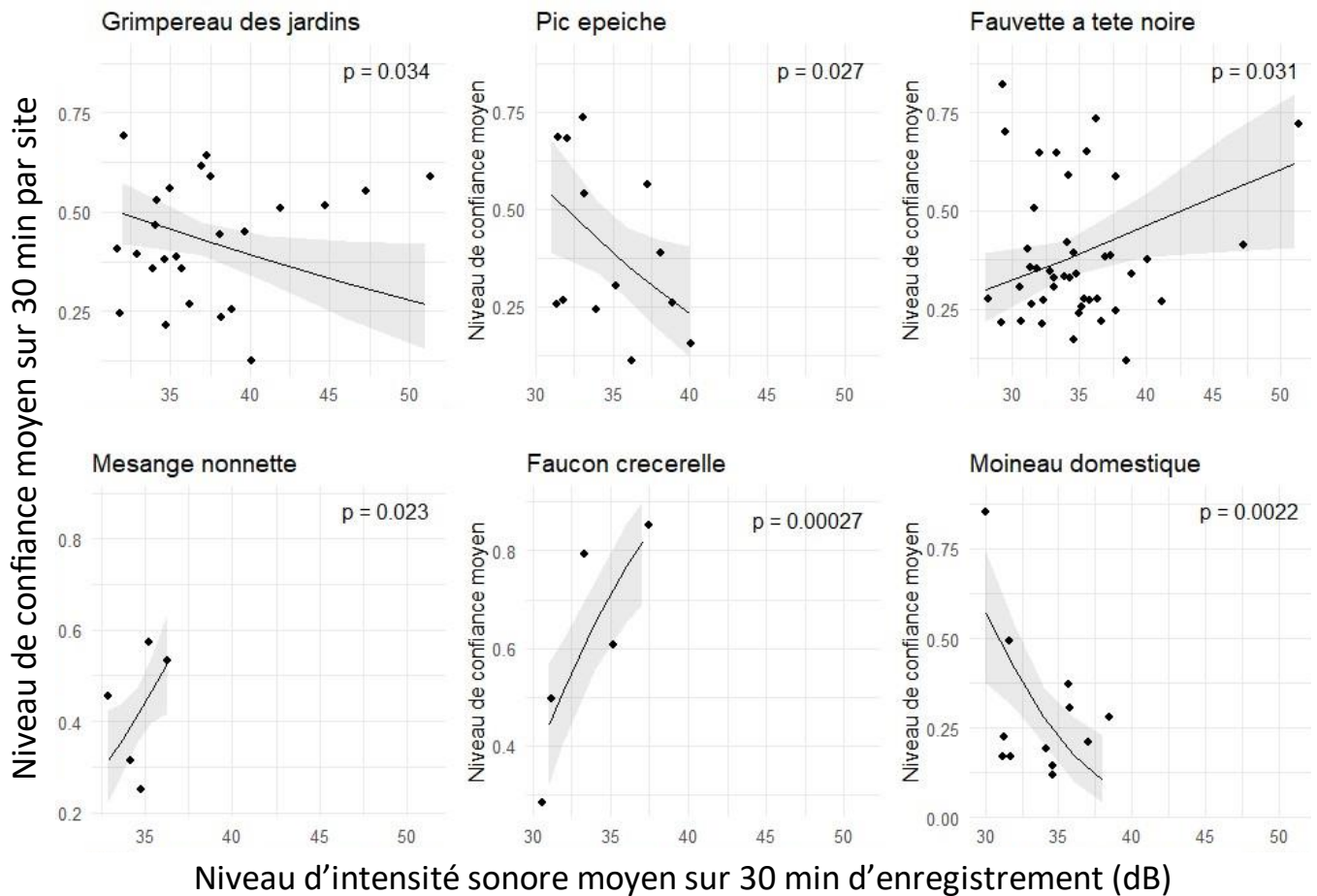


Figure 23 : Comportement du niveau de confiance moyen pour différentes espèces en fonction du niveau d'intensité sonore. Les données concernent 30 min d'enregistrement. Les espèces affichées sont celles pour lesquelles la relation s'est avérée significative

6.c.ii. Effet du NDSI sur le niveau de confiance moyen

Le grimpereau des jardins, le moineau domestique, le pic épeiche et le pigeon ramier voient leur niveau de confiance moyen diminuer pour des valeurs de NDSI tendant vers 1, i.e. des milieux sonores dominés par la composante biophonique (figure 24). Les données de niveau de confiance moyen du pic épeiche ne concernent pas des paysages sonores dominés par l'anthropophonie. Comme précédemment, les niveaux de confiance moyens associés à la fauvette noire ont un comportement différent vis-à-vis de la variable de pollution sonore par rapport aux autres espèces. En effet, ils augmentent de manière nette en passant d'un paysage sonore dominé par l'anthropophonie à un paysage sonore où prédominent les sons d'origine biologique (figure 24). De même que précédemment, peu de conclusions peuvent être tirées des données de la mésange nonnette et le martinet noir.

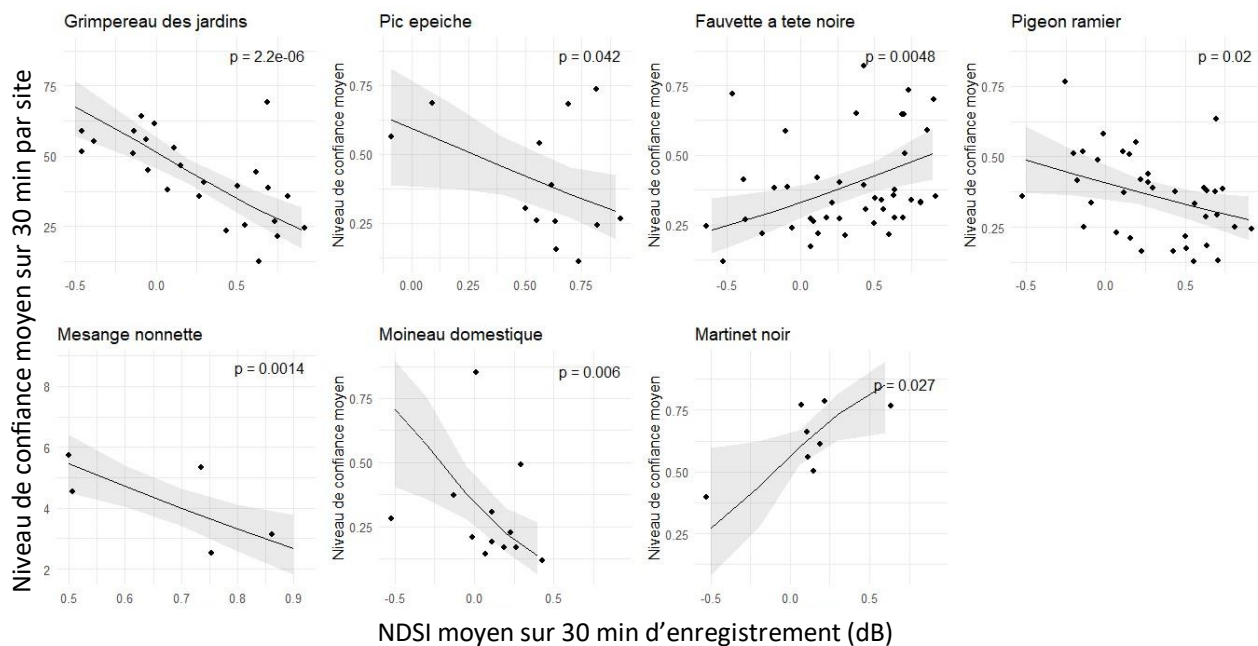


Figure 24 : Comportement du niveau de confiance moyen pour différentes espèces en fonction du NDSI. Les données concernent 30 min d'enregistrement. Les espèces affichées sont celles pour lesquelles la relation s'est avérée significative.

Finalement, l'effet de la pollution sonore n'est pas généralisable à toutes les espèces et environnements sonores et peut être contre-intuitif. D'une part, le niveau d'intensité sonore, i.e. si le site est bruyant, semble avoir un effet négatif sur la détection, se traduisant par une baisse dans la certitude des détections de BirdNET, bien que cet effet ne soit parfois discutable. D'autre part, le NDSI nous informe du contexte de chant des espèces pour lesquelles le modèle est significatif. Ainsi, la détection de certaines espèces (moineau domestique, grimpereau des jardins) s'améliore pour des paysages sonores dominés par l'anthropophonie. Ces espèces sont aussi présentes dans des milieux plus forestiers, mais ne parviennent pas à s'imposer dans des paysages sonores dominés par la biophonie, ce qui se retrouve dans leurs niveaux de confiance moyens qui sont alors plus faibles.

IV) DISCUSSION

1. Performance de la méthode automatisée

1.a. Bilan des résultats de la performance de BirdNET

L'objectif de cette étude étant d'évaluer les performances de la méthode automatisée dans la détection des communautés d'oiseaux sur les sites du suivi Evolville, elle applique une méthodologie peu commune dans la littérature considérant l'emploi de BirdNET à l'échelle d'une communauté et non d'une espèce, en appliquant un niveau de confiance unique pour filtrer les données issues de l'analyse des enregistrements sonores par l'algorithme (Funosas et al., 2024). Les études s'intéressant à l'analyse des performances de BirdNET ou autre algorithme de reconnaissance vocale des espèces considèrent souvent les niveaux de confiance à appliquer par espèce afin de maximiser la précision de la détection (Barré et al., 2019 ; Bota et al., 2023). L'inconvénient de cette méthode est la production d'une multitude d'indices de confiance à appliquer, et la difficulté à considérer plusieurs espèces dans une même étude. De plus, un consensus existe dans la littérature quant à l'irreproductibilité des filtres entre les sites et études pour une même espèce. Par exemple, dans son article de revue, Perez-Granados résume que le grand corbeau (*Corvus corax*) a obtenu 3 niveaux de confiance optimaux dans 3 études différentes : 0.29 (Cole et al., 2022), 0.66 (Kahl et al., 2020) et 0.94 (Sethi et al., 2021). Ces incertitudes incitent

néanmoins à l'exploration des performances de BirdNET avec des métriques similaires (notamment précision et sensibilité) afin d'améliorer la compréhension des performances de l'algorithme en fonction de l'espèce, du paysage, de l'environnement sonore et de l'enregistreur.

Finalement, le but poursuivi par l'étude doit guider le choix d'un niveau de confiance. S'il s'agit de détecter une espèce peu bavarde, rarement observée sur les sites d'étude, alors il convient de maximiser la probabilité de détecter cette espèce conservant les détections de BirdNET mais en les vérifiant manuellement. S'il s'agit, comme dans le cas de cette étude, de favoriser la précision à la sensibilité, il convient d'employer des filtres sélectifs. Certaines études préconisent des niveaux de confiance seuils : 0.5 (Wood et al., 2021 ; Pérez-Granados, 2023), 0.7 à 0.8 (Sethi et al., 2021).

Dans cette étude, nous n'avons identifié de manière certaine des valeurs de filtres que pour des durées de 30 minutes d'enregistrement, en confondant les paysages sonores. En maximisant la précision et limitant le nombre de faux négatifs, nous identifions la valeur de NC = 0.69 correspondant au maximum de F0.5-score. Pour ce filtre, la similarité entre les communautés de référence obtenues à travers la méthode semi-automatisée est de 0.4, soit une valeur assez faible, 45% des espèces présentes sont détectées, et la précision est de 0.82. Malgré l'absence de référence pour des enregistrements plus longs, on conclut à l'utilisation d'un filtre très sélectif (NC = 0.95) dans le but de favoriser la précision, que l'on espère au moins conserver au niveau précédemment décrit.

1.b. Les erreurs : les taxons concernés et raisons de ces erreurs

Certains taxons semblent plus affectés par les erreurs de la méthode. Premièrement, parmi les FN sont à compter les espèces produisant des cris et espèces spécialistes de milieux aquatiques. Parmi ces espèces, certains genres sont détectés correctement, mais pas l'espèce (cas des goélands notamment, dont l'espèce est également difficile à reconnaître sur le terrain), ce qui pousse à envisager un regroupement des détections d'espèces à l'échelle du genre pour certains taxons (Goélands, pigeons, pics) bien que la perte d'information puisse être problématique. Le problème de la détection de certaines espèces imitant les chants d'autres espèces (ex : geai des chênes) demeure également irrésolu (Kahl, 2020), bien que cela ne constitue pas un obstacle majeur à l'utilisation de BirdNET comme ces espèces ne sont pas fréquentes sur les relevés.

1.c. Importance de la liste d'espèces initiale

Le choix de la liste d'espèces servant de base de recherche de motifs à BirdNET doit faire l'objet d'une attention particulière. En effet, la liste utilisée dans cette étude (liste rouge des espèces nicheuses d'Alsace, allégée de quelques espèces impossibles à trouver au sein de l'EMS et enrichie d'espèces observées sur le terrain (ex : merle à plastron)) s'est avérée être encore trop large. Par exemple, cette liste a conduit à nombre d'identifications d'oiseaux nocturnes (chevêche d'Athéna, effraie des clochers), d'oiseaux d'eau (bihoreau gris, blongios nain) qui auraient facilement pu être évitées. Une liste d'espèces trop large affecte l'analyse globale des motifs par BirdNET notamment en contribuant à l'apparition de FP (Kahl, 2020).

Une proposition de liste d'espèce de référence pour cette étude serait de combiner les listes d'espèces obtenues sur les points d'écoute lors des 2 passages de 2024, ainsi que celles obtenues grâce à l'analyse manuelle d'enregistrements par l'expert. Dans le but d'ouvrir la porte aux détections d'espèces nouvelles, il serait intéressant de considérer la date et l'heure d'enregistrement visées et d'ajouter en conséquence à partir de la liste rouge des espèces d'Alsace les espèces dont la présence est plausible sur les sites de l'étude en tenant compte des paysages.

2. La pollution sonore est-elle problématique ?

Les résultats précédents ont montré que l'effet de la pollution sonore était propre aux différentes espèces. Des tendances semblent se dégager qu'il convient d'interpréter afin de comprendre dans quelle mesure les paysages sonores urbains sont propices aux analyses de BirdNET.

2.a. La taille des sites, une partie de l'explication ?

Les résultats sur les liens entre les variables de pollution sonore et les caractéristiques du paysage ont montré que les sites grands et avec un couvert végétalisé dominant ont un paysage sonore caractérisé par une biophonie dominante. A l'inverse, les sites petits et urbains, à proximité de routes sont les sites avec une anthropophonie dominante et plus bruyants. On peut penser que la taille plus réduite des sites urbains contribue grandement à la détectabilité de ces espèces pourtant sensibles aux niveaux d'intensité sonore élevés. C'est donc la configuration du site qui favorise la reconnaissance de ces espèces plutôt que leurs caractéristiques de chant, ce qui expliquerait que sur des sites plus grands, où d'autres oiseaux chantent, en milieu moins urbains, elles soient dominées vocalement par d'autres espèces.

Ainsi le NDSI est plutôt révélateur des conditions d'écoute que de l'effet de la pollution sonore en elle-même. Ces résultats tendent à montrer que l'effet du bruit urbain n'est pas source en soi d'une reconnaissance moins bonne de BirdNET comme les configurations des sites sont plutôt à l'œuvre. Les paysages sonores associés aux milieux urbains et moins urbains présentent chacun leurs défis et la taille du site et la proximité des oiseaux à l'appareil semble être un facteur clé de la méthode.

2.b. L'analyse visuelle des spectrogrammes pour différents paysages sonores offre des explications pour le comportement des niveaux de confiance associés aux détections d'espèces

Si l'on raisonne à l'échelle du spectrogramme, ces interrogations se traduisent par le recouvrement des signaux en termes de fréquence et leur intensité (niveau d'intensité sonore). 4 sites ont été sélectionnés, présentant un gradient de NDSI (moyenne de NDSI sur 30 min d'enregistrement). Les spectrogrammes des mêmes enregistrements de 5 min débutant à 6h39 pour des dates proches (entre le 6 et 8 juin 2024) ont été comparés.

Pour des milieux dont le NDSI est négatif, où l'anthropophonie domine, on observe qu'il y a un recouvrement fort de l'anthropophonie (bruits de circulation routière, son continu, sans forme spécifique, bande sonore généralement aux basses fréquences entre 0 et 3000 Hz sur les spectrogrammes), sur les chants d'oiseaux (figure 25). On observe néanmoins des motifs de chants distincts et de forte intensité (couleur tendant vers le jaune) lorsqu'il n'y a pas de recouvrement entre l'anthropophonie et la biophonie. Ainsi la situation semble binaire pour des sites où le paysage sonore est dominé par l'anthropophonie. Soit il n'y a pas de recouvrement entre anthropophonie et biophonie et la proximité des oiseaux à l'appareil permet de produire un motif net sur le spectrogramme. Soit il y a un recouvrement, plus ou moins prononcé et le motif de chant se fond dans le motif sonore lié à la circulation automobile (source de pollution anthropique sur ces enregistrements). Ainsi, soit BirdNET détecte un signal et grâce à sa netteté il y associe un niveau de confiance élevé, soit il ne détecte rien. Ceci corrobore les résultats précédents pour lesquels le degré de certitude de BirdNET associé à la détection de certaines espèces s'améliore en milieu urbain. Cependant, les espèces avec des fréquences de chant graves comme les corvidés sont potentiellement désavantagées sur des sites avec des composantes sonores anthropiques fortes. On peut donc supposer, sans affirmer, qu'en milieu urbain la précision pourrait être meilleure, i.e. le nombre de FP faibles, mais la sensibilité plus faible soit un nombre de FN élevé, notamment pour les espèces lointaines où avec des fréquences plutôt basses. La littérature montre que des déplacements des vocalisations d'oiseaux vers des fréquences de chant plus aigus ont été observés, de manière à améliorer leur distinction des sons anthropiques (Slabbekoorn &

Peet, 2003). Parmi les adaptations aux paysages urbains, on compte également un allongement de la durée des chants, ainsi que leur lenteur (Marín-Gómez & MacGregor-Fors, 2021).

Site 33 : NDSI = -0.38

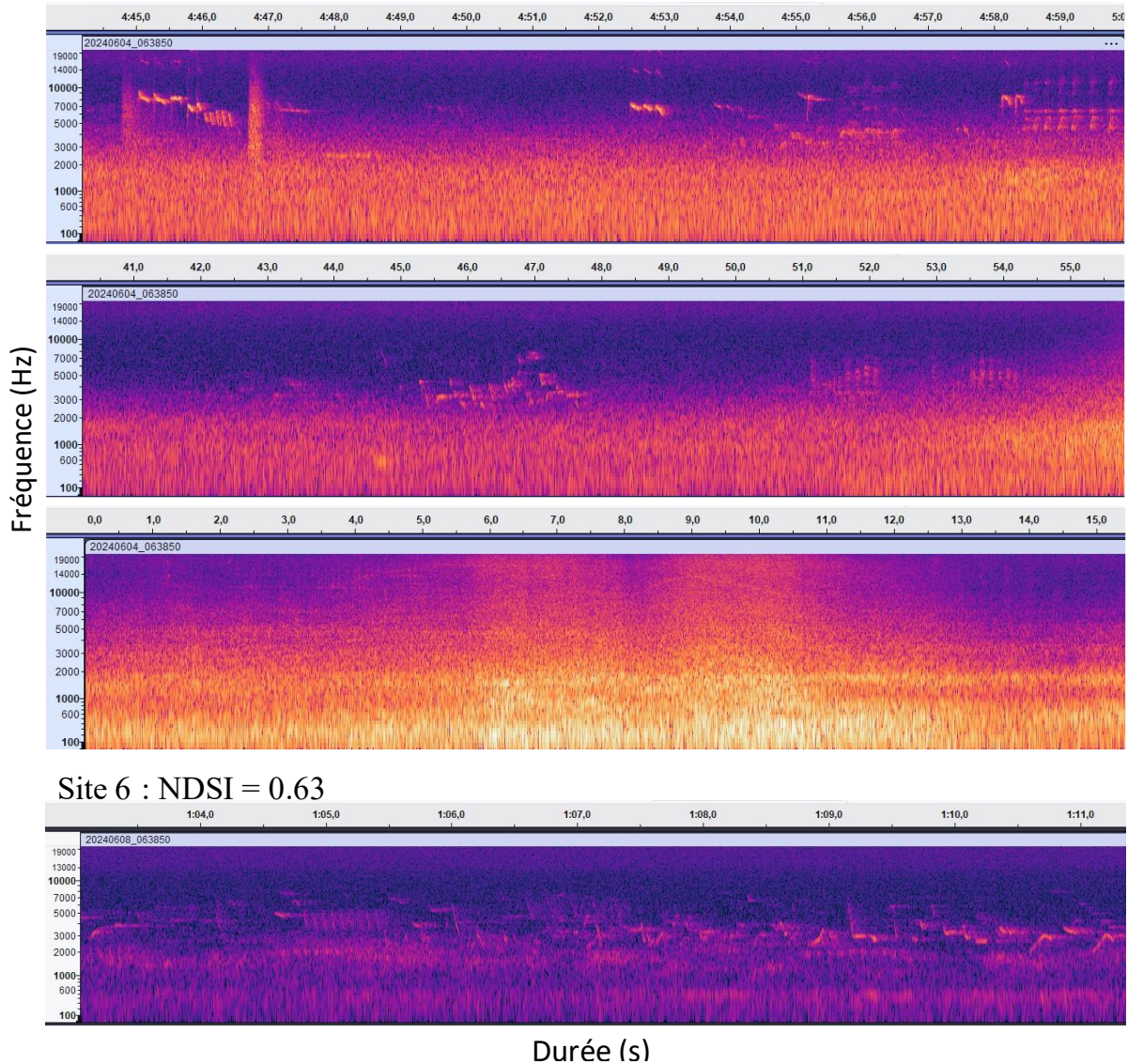


Figure 25 : Spectrogrammes obtenus pour différents sites de l'étude. Plus la couleur du motif tend vers le blanc, plus le niveau d'intensité sonore du signal est élevé. Les sites ont été choisis d'une manière à avoir des spectrogrammes pour des paysages sonores où l'anthropophonie domine (site 33) et où la biophonie domine (site 6). Les enregistrements sont pris au même horaire, le même mois à 4 jours d'écart (4 et 8 juin 2024). Le spectrogramme a été obtenu grâce au logiciel Audacity.

Pour des sites où la composante sonore anthropique est faible ou absente, le recouvrement des signaux n'est plus un problème. En conséquence, une multitude de motifs liés aux chants d'une diversité d'espèces plus grande apparaissent, plus ou moins intenses (figure 25). Le site 6 est un cas extrême, pour lequel le paysage sonore est celui d'une forêt. A la lecture du spectrogramme du site 6, BirdNET semble effectuer des reconnaissances sur tous les signaux, de différentes intensités, qui se mélangent. L'algorithme peine alors plus à distinguer les différentes espèces, ce qui expliquerait des niveaux de confiance associés aux détections plus faibles. On peut alors supposer que pour ce type de paysage

sonore le nombre de FP est un problème, alors que la sensibilité de l'algorithme l'est moins. La corrélation positive entre le *recall* et la biophonie a par ailleurs été démontrée (Funosas et al., 2024).

Précédemment, on a conclu à l'emploi nécessaire de niveaux de confiance élevés dans le cadre de l'utilisation d'enregistrements longs. L'analyse des spectrogrammes pousse à supposer que cette méthode est applicable sur les sites le long du gradient urbain-forestier. En effet, cela éliminerait nombre de FP pour des paysages sonores dominés par la biophonie et sauvegarderait les signaux d'intensité forte pour des paysages sonores où les sons anthropiques sont très présents. On peut émettre l'hypothèse que les deux facettes de la performance sont affectées différemment par le paysage sonore : la sensibilité est en jeu dans des paysages sonores anthropisés et la précision dans des paysages sonores biophoniques.

2.c. Récapitulatif et interprétation de l'effet de la pollution sonore sur la performance de BirdNET

Les analyses précédentes sur l'effet de la pollution sonore ont montré l'effet négatif du niveau sonore sur les identifications lorsque celui-ci est élevé. Les paysages sonores bruyants se sont également avérés être les paysages sonores urbains. Or, contre-intuitivement, le niveau de confiance moyen de certaines espèces s'est avéré meilleur dans des paysages sonores dominés par l'anthropophonie. Ces espèces (Grimpereau des jardins, pigeon ramier, moineau domestique) sont trouvées fréquemment en milieu urbain. La petite taille des sites en milieu urbain contribue à la dichotomie entre les chants qui se distinguent de la bande sonore du bruit anthropique de manière claire et qui sont alors détectés avec une certitude élevée par BirdNET, et les chants qui se fondent dans la bande son du bruit anthropique et qui ne sont pas détectés. Certaines études décrivent l'adaptation du chant des espèces en fonction de l'anthropophonie, se traduisant par une augmentation des fréquences de chant et un changement de phénologie, menant à une distinction accrue de l'anthropophonie et les chants et donc des détections de BirdNET plus précises (Slabbekoorn, 2013 ; Lohr et al., 2003). Cependant, ces espèces sont moins bien détectées dans des milieux dominés par la biophonie. On peut penser que la distance à l'enregistreur constitue un obstacle, mais aussi que les chants de ces espèces ne sont pas dominants.

Le cas des espèces associées aux milieux forestiers semble différent. Dans l'étude, la fauvette à tête noire est mieux détectée dans des environnements sonores dominés par la biophonie et il est difficile de conclure à sa sensibilité au niveau d'intensité sonore. On peut penser que les chants de cet oiseau sont dominants dans un chorus, ce qui se traduit par un niveau d'intensité sonore du chant plus élevé et un motif plus intense sur le spectrogramme, facilitant la reconnaissance de BirdNET. Cette espèce étant plus sensible aux environnements urbains, on peut penser qu'elle y chante moins (Marín-Gómez & MacGregor-Fors, 2021) et qu'il existe éventuellement un recouvrement par l'anthropophonie (Lohr et al., 2003). Toujours pour les espèces forestières, un deuxième cas est représenté par nos résultats. Le pic épeiche est trouvé dans notre relevé dans des paysages sonores biophoniques et relativement silencieux. La certitude associée à la détection de cette espèce diminue lorsque le NDSI augmente, témoignant d'un chorus plus puissant dans le cas de notre relevé. On a alors le cas d'une espèce dont le motif de chant perd en intelligibilité sur des sites grands et dont l'environnement sonore est chargé en chants d'autres espèces. Cette espèce émettant des cris, cela pourrait contribuer à la difficulté de sa reconnaissance et son éventuelle confusion avec des espèces du même genre.

Les paysages sonores urbains sont donc propices à la reconnaissance des espèces par BirdNET grâce à la proximité de l'oiseau et aux paramètres intrinsèques du chant pour les espèces urbanophiles. Dans des milieux plus forestiers, où prédominent les chants d'oiseaux, la puissance du chant va être déterminante pour la détection de BirdNET et l'abondance de motifs peu clairs va certainement nuire à la certitude accordée à chaque détection et produire un nombre important de FP. Néanmoins ces interprétations restent à confirmer par des analyses statistiques.

3. La question des espèces rares

Notre étude ne permet pas de conclure quant à la capacité de BirdNET à détecter les espèces rares en fonction des différents paysages sonores. L'étude de l'effet de la pollution sonore a intégré quelques espèces plus rares : Faucon crécerelle, Lorient d'Europe, Mésange nonnette, Pouillot véloce, Verdier d'Europe, pour lesquelles aucun lien significatif n'a pu être trouvé.

La détection des espèces rares dépend de la probabilité de leur présence sur les sites lors de la pose de l'appareil. De plus, il a été montré que les oiseaux dans des paysages sonores urbains chantent moins et si l'on considère des espèces rares, la probabilité de capter leur présence est d'autant plus réduite. Quant au chant, la littérature montre que les espèces urbanophiles tendent à l'adapter aux paysages sonores anthropisés. La flexibilité comportementale dont fait partie la plasticité acoustique est un élément clé de l'adaptation des espèces au milieu urbain (Slabbekoorn, 2013). On peut alors penser que des espèces rares et urbanophobes, ne bénéficient pas des adaptations de chant des espèces urbanophiles et seraient ainsi moins détectables. Il est alors possible d'envisager que la détection des espèces rares en milieu urbain dépende de la contingence de la non-superposition des bandes sonores liées à l'anthropophonie et aux chants, ce qui implique que l'espèce doit chanter à des fréquences au-dessus de 2000 à 3000 Hz, à un instant plus silencieux. Ainsi on voit apparaître l'hypothèse que la méthode est plus efficace pour des espèces plus communes (Funosas et al., 2024), bien que l'étude ne permette pas de l'affirmer. Finalement, il est question de savoir si les paramètres propres aux sites urbains (proximité de l'animal à l'enregistreur grâce à la petite taille des sites, densité d'oiseaux plus faible (Slabbekoorn, 2013) réduisant le chevauchement entre les chants et donc augmentant la clarté du signal) pourraient compenser le déficit de détection lié à la pollution sonore.

4. BirdNET : Le rôle des experts naturalistes

Une bonne gestion des aléas dans l'utilisation de BirdNET (paramétrage, niveau de confiance, liste initiale d'espèces, filtre sur le nombre de points de contact minimal pour une espèce par site) (Funosas et al. 2024) assure une précision suffisante pour employer les données produites. La longueur des enregistrements est cruciale d'un point de vue probabiliste pour capturer le plus d'espèces possibles de la communauté. Cependant, notre étude souligne l'impossibilité d'utiliser la méthode automatisée seule. En fonction des objectifs de l'étude, des points d'écoute ou bien des vérifications manuelles des enregistrements doivent être faits. Surtout, il convient d'avoir une idée claire des espèces trouvées sur la région d'étude, d'établir une liste de référence à partir d'observations vérifiées par des experts. Le choix d'ouvrir la liste à des espèces potentiellement sur les sites revient à l'auteur de l'étude et doit s'effectuer précautionneusement. Il est important de rappeler que l'observation directe sur le terrain permet de détecter des espèces différentes de BirdNET, notamment les espèces d'eau, émettant des cris ou plus silencieuses.

5. BirdNET : mode d'emploi sur les sites Evolville

Pour résumer, plusieurs aspects clés sont à prendre en compte dans l'utilisation de BirdNET pour améliorer la fiabilité de la méthode (figure 26). Le paramétrage initial de l'application est un premier élément et les valeurs de sensitivity et recouvrement ont une influence sur la performance de l'algorithme (Funosas et al., 2024). La liste d'espèces de référence de BirdNET est essentielle comme elle va favoriser ou non l'apparition de Faux Positifs dans les relevés. La durée enregistrée va déterminer l'accumulation de données et l'augmentation du nombre de détections, avec notamment des niveaux de confiance plus élevés. Enfin le paysage sonore est à prendre en compte dans la construction d'un protocole d'utilisation de BirdNET pour savoir quelles vont être les natures des biais du relevé : sensibilité exacerbée de l'algorithme à l'origine de Faux Positifs potentiels ou camouflage d'une partie des chants par l'anthropophonie.

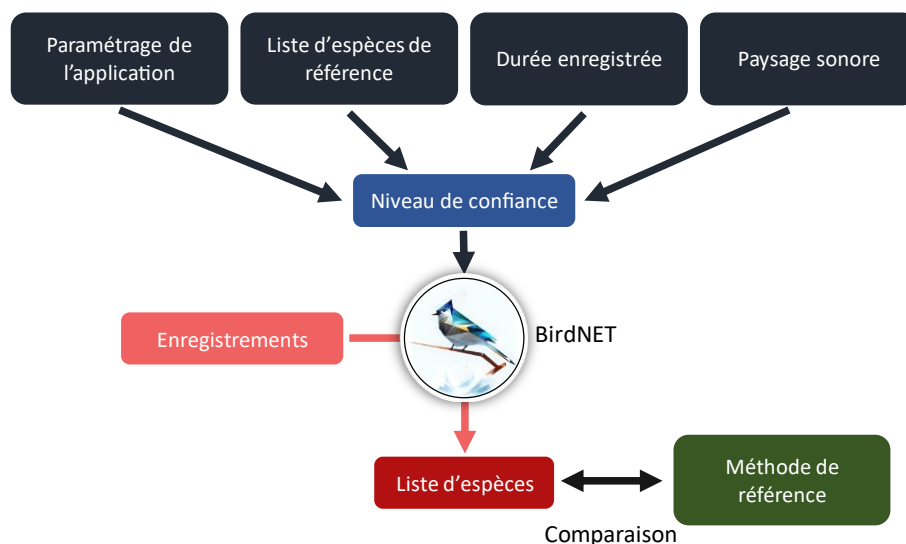


Figure 26 : Résumé des étapes clés dans l'utilisation de BirdNET.

Tous ces paramètres ajustables vont avoir une influence sur le niveau de confiance à choisir pour filtrer les données de BirdNET. La liste d'espèces produite par BirdNET devra être comparée aux analyses réalisées avec un point d'écoute ou l'écoute de l'enregistrement par l'expert, notamment en cas de doute sur la présence d'une espèce.

6. Limites et perspectives de l'étude

L'étude souffre de limites supplémentaires à celles déjà abordées et qu'il convient de présenter. Concernant les enregistrements, la variabilité des appareils telle que la qualité du microphone et la présence d'un bruit de fond est réelle mais n'a pas été étudiée. L'effet de ces paramètres sur la performance de BirdNET ainsi que sur les caractéristiques de pollution sonore des sites est inconnu.

En ce qui concerne l'étude de la pollution sonore à partir de données produites par l'analyse des enregistrements par R, la méthode souffre du manque d'informations sur la véracité de cette analyse et du peu de données produites sur le terrain. Le protocole de relevé de données de pollution sonore sur le terrain est globalement à revoir, avec la nécessité de mettre en place un plan d'échantillonnage sur l'ensemble des sites, combinant au moins deux appareils performants de relevé de la pollution sonore. Il serait également idéal de réaliser ces relevés simultanément sur les sites. Sur le terrain, le niveau d'intensité sonore maximal était relevé, alors que l'on peut s'interroger sur la contingence d'un son ponctuel puissant qui viendrait fausser cette carte d'identité sonore du site.

Le choix de métriques pour décrire la pollution sonore est sujet à controverse comme l'écologie du son est une discipline récente. De multiples indicateurs ont été produits afin d'automatiser l'analyse écologique d'enregistrements acoustiques, comme l'Acoustic Complexity Index (ACI) (Pieretti et al., 2011), Acoustic Diversity Index (ADI) (Villanueva-Rivera et al., 2011), Bioacoustic Index (BI) (Boelman et al., 2007), etc., avec des objectifs divers et complémentaires. Cependant, la finesse de ces indicateurs reste méconnue et certaines études prônent la plus grande prudence quant à leur emploi comme ils ne sont pas stables à travers les différents paysages sonores (Fairbrass et al., 2017).

L'approche multispécifique de l'étude constitue un aspect central de la méthode qu'il est possible de questionner comme elle ignore les biais liés à chaque espèce. Le choix de cette méthode s'est fait avec l'objectif pratique de vérifier sa performance pour sa mise en place sur les sites de l'observatoire de la biodiversité urbaine Evolville. Il ne s'agit pas de maximiser la précision par espèce, de façon que le plus

grand nombre de chants soient reconnus correctement, mais plutôt qu'un nombre défini de chants soient reconnus correctement pour enregistrer la présence de l'espèce. Le *recall* n'intervient que dans la mesure où l'on cherche à détecter l'espèce, mais la maximisation de la sensibilité n'est pas un objectif. En employant un filtre sévère, des espèces sont perdues sur tous les sites de l'étude. A défaut d'une référence, pour des enregistrements de matinées complètes, il est impossible de savoir quelles espèces ont été ignorées alors qu'elles étaient présentes et il est certain que des Faux Positifs seront aussi présents dans le relevé. En comparant les analyses de BirdNET sur une matinée complète avec les listes des espèces obtenues avec la méthode semi-automatisée sur 30 min, on peut néanmoins conclure qu'avec $NC = 0.95$, le relevé perdrait par exemple le bruant des roseaux ($NC_{max} = 0.67$), le pic cendré ($NC_{max} = 0.93$), l'hypolaïs polyglotte ($NC_{max} = 0.75$), le tarier pâtre ($NC_{max} = 0.35$), sans savoir cependant si ces espèces étaient détectées sur les bons sites avec ces points de contact et donc si l'on perd un Vrai Positif. Or ces espèces sont moins fréquentes sur les sites ce qui pourrait indiquer que la méthode est moins favorable à la détection d'espèces plus rares.

V) CONCLUSION

BirdNET s'est imposé comme l'algorithme de référence en matière de reconnaissance des chants d'oiseaux et ses performances restent inégalées. Sa diffusion est facilitée auprès du grand public grâce à l'application mobile et son usage est plébiscité par la communauté scientifique qui y voit une méthode avantageuse d'acquisition des données. Les limites de cet outil restent toutefois inconnues. Cette étude s'inscrit dans la démarche d'étudier le lien entre les analyses de BirdNET et le paysage sonore et questionne notamment l'utilisation de BirdNET pour décrire les communautés d'oiseaux présentes selon un gradient urbain-forestier selon une méthode unifiée.

Le long du gradient urbain-forestier, la précision de BirdNET s'est avérée être bonne (précision = 0.82), mais en contrepartie, plus de la moitié des espèces ne sont pas détectées (*recall* = 0.45) pour une durée de 30 min d'enregistrement analysées, pour un niveau de confiance valant 0.69. L'emploi de BirdNET pour analyser des enregistrements à durées longues (matinées entières) est une méthode valide avec des filtres sélectifs (niveau de confiance = 0.95) sur les identifications pour limiter le nombre de faux positifs qui est un aléa majeur dans l'étude des communautés avec cet algorithme.

D'autre part, l'étude des liens entre paysages sonores, caractéristiques des sites et performance de l'algorithme a permis de mettre en évidence que l'organisation des sites urbains et forestiers ainsi que des communautés d'oiseaux s'y trouvant influençaient de manière différente la détection de BirdNET, qui s'avère utilisable le long du gradient urbain-forestier. Un résultat important de l'étude est la bonne performance de l'algorithme pour des espèces urbanophiles notamment, se traduisant par une certitude plus grande accordée aux détections des TP, ce qui permet d'augmenter la sévérité du filtre pour éliminer les FP tout en conservant les TP. On peut relier cette bonne performance en milieu urbain à la proximité des espèces à l'enregistreur combinée à la densité plus faible d'individus contribuant à la clarté des chants sur les spectrogrammes. Sur les sites forestiers, l'absence d'overlap entre la biophonie et anthropophonie contribue à une sensibilité exacerbée de l'algorithme qui détecte mieux les espèces ayant un chant puissant, se distinguant du chœur. Finalement, le niveau d'intensité sonore est défavorable à l'analyse de BirdNET, mais ne suffit pas à expliquer la performance de l'algorithme sur les sites. Ceci requiert l'absence d'overlap entre les motifs de chants de différentes espèces entre elles et avec l'anthropophonie, ainsi qu'une puissance de chant suffisante pour imprimer le motif sur le spectrogramme.

Les résultats de cette étude sont optimistes quant à l'usage de BirdNET sur des enregistrements en contexte urbain, bien qu'approfondir l'analyse de l'effet de la pollution sonore soit à envisager. En effet, cette étude a permis de soulever la question du lien entre la spécialisation écologique de l'oiseau, ses capacités vocales (notamment la plasticité) et le paysage sonore dans lequel elle est entendue.

BIBLIOGRAPHIE :

- Aumond, P., Lavandier, C., Ribeiro, C., Boix, E. G., Kambona, K., D'Hondt, E., & Delaitre, P. (2017). A study of the accuracy of mobile technology for measuring urban noise pollution in large scale participatory sensing campaigns. *Applied Acoustics*, 117, 219-226. <https://doi.org/10.1016/j.apacoust.2016.07.011>
- Barré, K., Le Viol, I., Julliard, R., Pauwels, J., Newson, S. E., Julien, J.-F., Claireau, F., Kerbiriou, C., & Bas, Y. (2019). Accounting for automated identification errors in acoustic surveys. *Methods in Ecology and Evolution*, 10(8), 1171-1188. <https://doi.org/10.1111/2041-210X.13198>
- Besson, M., Alison, J., Bjerger, K., Gorochowski, T. E., Høye, T. T., Jucker, T., Mann, H. M. R., & Clements, C. F. (2022). Towards the fully automated monitoring of ecological communities. *Ecology Letters*, 25(12), 2753-2775. <https://doi.org/10.1111/ele.14123>
- Bonthoux, S., Brun, M., Di Pietro, F., Greulich, S., & Bouché-Pillon, S. (2014). How can wastelands promote biodiversity in cities? A review. *Landscape and Urban Planning*, 132, 79-88. <https://doi.org/10.1016/j.landurbplan.2014.08.010>
- Bota, G., Manzano-Rubio, R., Catalán, L., Gómez-Catasús, J., & Pérez-Granados, C. (2023). Hearing to the Unseen : AudioMoth and BirdNET as a Cheap and Easy Method for Monitoring Cryptic Bird Species. *Sensors*, 23(16), Article 16. <https://doi.org/10.3390/s23167176>
- Bradfer-Lawrence, T., Desjonqueres, C., Eldridge, A., Johnston, A., & Metcalf, O. (2023). Using acoustic indices in ecology : Guidance on study design, analyses and interpretation. *Methods in Ecology and Evolution*, 14(9), 2192-2204. <https://doi.org/10.1111/2041-210X.14194>
- Brooker, S. A., Stephens, P. A., Whittingham, M. J., & Willis, S. G. (2020). Automated detection and classification of birdsong : An ensemble approach. *Ecological Indicators*, 117, 106609. <https://doi.org/10.1016/j.ecolind.2020.106609>
- Brumm, H., & Slabbekoorn, H. (2005). Acoustic Communication in Noise. In *Advances in the Study of Behavior* (Vol. 35, p. 151-209). Academic Press. [https://doi.org/10.1016/S0065-3454\(05\)35004-2](https://doi.org/10.1016/S0065-3454(05)35004-2)
- Clark, M. L., Salas, L., Baligar, S., Quinn, C. A., Snyder, R. L., Leland, D., Schackwitz, W., Goetz, S. J., & Newsam, S. (2023). The effect of soundscape composition on bird vocalization classification in a citizen science biodiversity monitoring project. *Ecological Informatics*, 75, 102065. <https://doi.org/10.1016/j.ecoinf.2023.102065>
- Cole, J. S., Michel, N. L., Emerson, S. A., & Siegel, R. B. (2022). Automated bird sound classifications of long-duration recordings produce occupancy model outputs similar to manually annotated data. *Ornithological Applications*, 124(2), duac003. <https://doi.org/10.1093/ornithapp/duac003>
- Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in R. *Journal of Statistical Software*. <https://www.jstatsoft.org/article/view/v034i02>
- Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning - ICML '06*, 233-240. the 23rd international conference. <https://doi.org/10.1145/1143844.1143874>
- Fairbrass, A. J., Rennert, P., Williams, C., Titheridge, H., & Jones, K. E. (2017). Biases of acoustic indices measuring biodiversity in urban areas. *Ecological Indicators*, 83, 169-177. <https://doi.org/10.1016/j.ecolind.2017.07.064>
- Flach, P. A., & Kull, M. (2015, décembre 7). *Precision-Recall-Gain Curves : PR Analysis Done Right*. Neural Information Processing Systems. <https://www.semanticscholar.org/paper/Precision-Recall-Gain-Curves%3A-PR-Analysis-Done-Flach-Kull/f50f9c08056c50898e8ef1b93a14ae85d525bbd2>

- Fraixedas, S., Lindén, A., Piha, M., Cabeza, M., Gregory, R., & Lehtikoinen, A. (2020). A state-of-the-art review on birds as indicators of biodiversity : Advances, challenges, and future directions. *Ecological Indicators*, 118, 106728. <https://doi.org/10.1016/j.ecolind.2020.106728>
- Fuller, S., Axel, A. C., Tucker, D., & Gage, S. H. (2015). Connecting soundscape to landscape : Which acoustic index best describes landscape configuration? *Ecological Indicators*, 58, 207-215. <https://doi.org/10.1016/j.ecolind.2015.05.057>
- Funosas, D., Barbaro, L., Schillé, L., Elger, A., Castagneyrol, B., & Cauchoix, M. (2024). Assessing the potential of BirdNET to infer European bird communities from large-scale ecoacoustic data. *Ecological Indicators*, 164, 112146. <https://doi.org/10.1016/j.ecolind.2024.112146>
- Gerber, L. R., & Iacona, G. D. (2024). Aligning data with decisions to address the biodiversity crisis. *PLOS Biology*, 22(6), e3002683. <https://doi.org/10.1371/journal.pbio.3002683>
- Goëau, H., Glotin, H., Vellinga, W.-P., Planqué, R., Rauber, A., & Joly, A. (2014). *LifeCLEF Bird Identification Task 2014*.
- Kahl, S. (2020). *Identifying birds by sound : Large-scale acoustic event recognition for avian activity monitoring*. Universitätsverlag Chemnitz.
- Kahl, S. (2024, septembre 10). *Kahst/BirdNET-Lite*. <https://github.com/kahst/BirdNET-Lite> (Édition originale 2020)
- Kahl, S., Wood, C. M., Eibl, M., & Klinck, H. (2021). BirdNET : A deep learning solution for avian diversity monitoring. *Ecological Informatics*, 61, 101236. <https://doi.org/10.1016/j.ecoinf.2021.101236>
- Kasten, E. P., Gage, S. H., Fox, J., & Joo, W. (2012). The remote environmental assessment laboratory's acoustic library : An archive for studying soundscape ecology. *Ecological Informatics*, 12, 50-67. <https://doi.org/10.1016/j.ecoinf.2012.08.001>
- Keilwagen, J., Grosse, I., & Grau, J. (2014). Area under Precision-Recall Curves for Weighted and Unweighted Data. *PLoS ONE*, 9(3), e92209. <https://doi.org/10.1371/journal.pone.0092209>
- Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Căsar, C., Coen, M., DeRuiter, S. L., Doyle, L., Edelman, S., Ferrer-i-Cancho, R., Freeberg, T. M., Garland, E. C., Gustison, M., Harley, H. E., ... Zamora-Gutierrez, V. (2016). Acoustic sequences in non-human animals : A tutorial review and prospectus. *Biological Reviews*, 91(1), 13-52. <https://doi.org/10.1111/brev.12160>
- Koskimies, P. (1989). *Scopus—Document details—Birds as a tool in environmental monitoring*. <https://www.scopus.com/record/display.uri?eid=2-s2.0-0024932577&origin=resultslist&sort=plf-f&src=s&sid=0579c4336769f6ee13bbd271771e5f88&sot=b&sdt=cl&cluster=scopubyr%2C%221989%22%2Ct&s=TITLE-ABS-KEY%28Birds+as+a+tool+in+environmental+monitoring%29&sl=112&sessionSearchId=0579c4336769f6ee13bbd271771e5f88&relpos=0>
- Leach, E. C., Burwell, C. J., Ashton, L. A., Jones, D. N., & Kitching, R. L. (2016). Comparison of point counts and automated acoustic monitoring : Detecting birds in a rainforest biodiversity survey. *Emu - Austral Ornithology*, 116(3), 305-309. <https://doi.org/10.1071/MU15097>
- Lewandowski, A. S., Noss, R. F., & Parsons, D. R. (2010). The Effectiveness of Surrogate Taxa for the Representation of Biodiversity. *Conservation Biology*, 24(5), 1367-1377. <https://doi.org/10.1111/j.1523-1739.2010.01513.x>

- Lohr, B., Wright, T. F., & Dooling, R. J. (2003). Detection and discrimination of natural calls in masking noise by birds : Estimating the active space of a signal. *Animal Behaviour*, 65(4), 763-777. <https://doi.org/10.1006/anbe.2003.2093>
- Marcon, E. (2015). *Mesures de la Biodiversité* [Lecture, AgroParisTech]. <https://agroparistech.hal.science/cel-01205813>
- Marín-Gómez, O. H., & MacGregor-Fors, I. (2021). A global synthesis of the impacts of urbanization on bird dawn choruses. *Ibis*, 163(4), 1133-1154. <https://doi.org/10.1111/ibi.12949>
- McGinn, K., Kahl, S., Peery, M. Z., Klinck, H., & Wood, C. M. (2023). Feature embeddings from the BirdNET algorithm provide insights into avian ecology. *Ecological Informatics*, 74, 101995. <https://doi.org/10.1016/j.ecoinf.2023.101995>
- More than half of world's population now living in urban areas, UN survey finds* | UN News. (2014, juillet 10). <https://news.un.org/en/story/2014/07/472752>
- Muratet, A. (2006). *Diversité végétale en milieu urbain* [Université Pierre et Marie Curie ; Muséum National d'Histoire Naturelle]. <https://digitale.cbnbl.org/documents/THE351.pdf>
- Muratet, A., Hardion, L., Sotillo, A., Fujiki, K., & Chanez, E. (2024). *Plant responses to urban gradients : Extinction, plasticity, adaptation* [Jeu de données]. Zenodo. <https://doi.org/10.5281/zenodo.13757647>
- Muratet, A., Muratet, M., Pellaton, M., Brun, M., Baude, M., Shwartz, A., & Fontaine, C. (2021). Wasteland, a Refuge for Biodiversity, for Humanity. In F. Di Pietro & A. Robert (Éds.), *Urban Wastelands : A Form of Urban Nature?* (p. 95-120). Springer International Publishing. https://doi.org/10.1007/978-3-030-74882-1_5
- Napier, T., Ahn, E., Allen-Ankins, S., Schwarzkopf, L., & Lee, I. (2024). Advancements in preprocessing, detection and classification techniques for ecoacoustic data : A comprehensive review for large-scale Passive Acoustic Monitoring. *Expert Systems with Applications*, 252, 124220. <https://doi.org/10.1016/j.eswa.2024.124220>
- Niemelä, J. (2011). *Urban Ecology : Patterns, Processes, and Applications* (J. H. Breuste, T. Elmqvist, G. Guntenspergen, P. James, & N. E. McIntyre, Éd.). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199563562.001.0001>
- Nordt, A., & Klenke, R. (2013). Sleepless in Town – Drivers of the Temporal Shift in Dawn Song in Urban European Blackbirds. *PLoS ONE*, 8(8), e71476. <https://doi.org/10.1371/journal.pone.0071476>
- Pérez-Granados, C. (2023). BirdNET : Applications, performance, pitfalls and future opportunities. *Ibis*, 165(3), 1068-1075. <https://doi.org/10.1111/ibi.13193>
- Pérez-Granados, C., & Traba, J. (2021). Estimating bird density using passive acoustic monitoring : A review of methods and suggestions for further research. *Ibis*, 163(3), 765-783. <https://doi.org/10.1111/ibi.12944>
- Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., Gage, S. H., & Pieretti, N. (2011). Soundscape Ecology : The Science of Sound in the Landscape. *BioScience*, 61(3), 203-216. <https://doi.org/10.1525/bio.2011.61.3.6>
- Pijanowski, L. J. V.-R. and B. C. (2018). *soundecology : Soundscape Ecology* (Version 1.3.3) [Logiciel]. <https://cran.r-project.org/web/packages/soundecology/index.html>
- Présentation de l'Eurométropole*. (s. d.). Consulté 23 septembre 2024, à l'adresse <https://www.strasbourg.eu/presentation-eurometropole>
- Priyadarshani, N., Castro, I., & Marsland, S. (2018). The impact of environmental factors in birdsong acquisition using automated recorders. *Ecology and Evolution*, 8(10), 5016-5033. <https://doi.org/10.1002/ece3.3889>

- Ross, S. R. P.-J., O'Connell, D. P., Deichmann, J. L., Desjonquères, C., Gasc, A., Phillips, J. N., Sethi, S. S., Wood, C. M., & Burivalova, Z. (2023). Passive acoustic monitoring provides a fresh perspective on fundamental ecological questions. *Functional Ecology*, 37(4), 959-975. <https://doi.org/10.1111/1365-2435.14275>
- Schmidhuber, J. (2022). *Annotated History of Modern AI and Deep Learning* (arXiv:2212.11279). arXiv. <https://doi.org/10.48550/arXiv.2212.11279>
- Sethi, S. S., Fossøy, F., Cretois, B., & Rosten, C. M. (2021). Management relevant applications of acoustic monitoring for Norwegian nature – The Sound of Norway. In 31. Norsk institutt for naturforskning (NINA). <https://brage.nina.no/nina-xmlui/handle/11250/2832294>
- Slabbekoorn, H. (2013). Songs of the city : Noise-dependent spectral plasticity in the acoustic phenotype of urban birds. *Animal Behaviour*, 85(5), 1089-1099. <https://doi.org/10.1016/j.anbehav.2013.01.021>
- Slabbekoorn, H., & Peet, M. (2003). Birds sing at a higher pitch in urban noise. *Nature*, 424(6946), 267-267. <https://doi.org/10.1038/424267a>
- Somerfield, P. J., Clarke, K. R., & Gorley, R. N. (2021). A generalised analysis of similarities (ANOSIM) statistic for designs with ordered factors. *Austral Ecology*, 46(6), 901-910. <https://doi.org/10.1111/aec.13043>
- Song Meter SM4 Wildlife Audio Recorder*. (s. d.). Wildlife Acoustics. Consulté 3 septembre 2024, à l'adresse <https://www.wildlifeacoustics.com/products/song-meter-sm4>
- Stastny, J., Munk, M., & Juranek, L. (2018). Automatic bird species recognition based on birds vocalization. *EURASIP Journal on Audio, Speech, and Music Processing*, 2018(1), 19. <https://doi.org/10.1186/s13636-018-0143-7>
- STOC - LPO (Ligue pour la Protection des Oiseaux). (s. d.). Consulté 2 septembre 2024, à l'adresse <https://www.lpo.fr/la-lpo-en-actions/connaissance-des-especes-sauvages/suivis-ornithologiques/oiseaux-communs/stoc>
- 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(4), 772-781. <https://doi.org/10.3813/AAA.918757>
- Sugai, L. S. M., Silva, T. S. F., Ribeiro, J. W., & Llusia, D. (2019). Terrestrial Passive Acoustic Monitoring : Review and Perspectives. *BioScience*, 69(1), 15-25. <https://doi.org/10.1093/biosci/biy147>
- Toenies, M., & Rich, L. (2021). Advancing bird survey efforts through novel recorder technology and automated species identification. *California Fish and Wildlife Journal*, 107(2), 56-70. <https://doi.org/10.51492/cfwj.107.5>
- Torgo, L., & Ribeiro, R. (2009). Precision and Recall for Regression. In J. Gama, V. S. Costa, A. M. Jorge, & P. B. Brazdil (Éds.), *Discovery Science* (p. 332-346). Springer. https://doi.org/10.1007/978-3-642-04747-3_26
- Wood, C. M., & Kahl, S. (2024a). Guidelines for appropriate use of BirdNET scores and other detector outputs. *Journal of Ornithology*. <https://doi.org/10.1007/s10336-024-02144-5>
- Xie, J., Zhong, Y., Zhang, J., Liu, S., Ding, C., & Triantafyllopoulos, A. (2023). A review of automatic recognition technology for bird vocalizations in the deep learning era. *Ecological Informatics*, 73, 101927. <https://doi.org/10.1016/j.ecoinf.2022.101927>
- Zéro artificialisation nette ZAN quelle mise en oeuvre de la loi de 2023 | *vie-publique.fr*. (2024, avril 18). <https://www.vie-publique.fr/en-bref/293834-zero-artificialisation-nette-zan-quelle-mise-en-oeuvre-de-la-loi-de-2023>

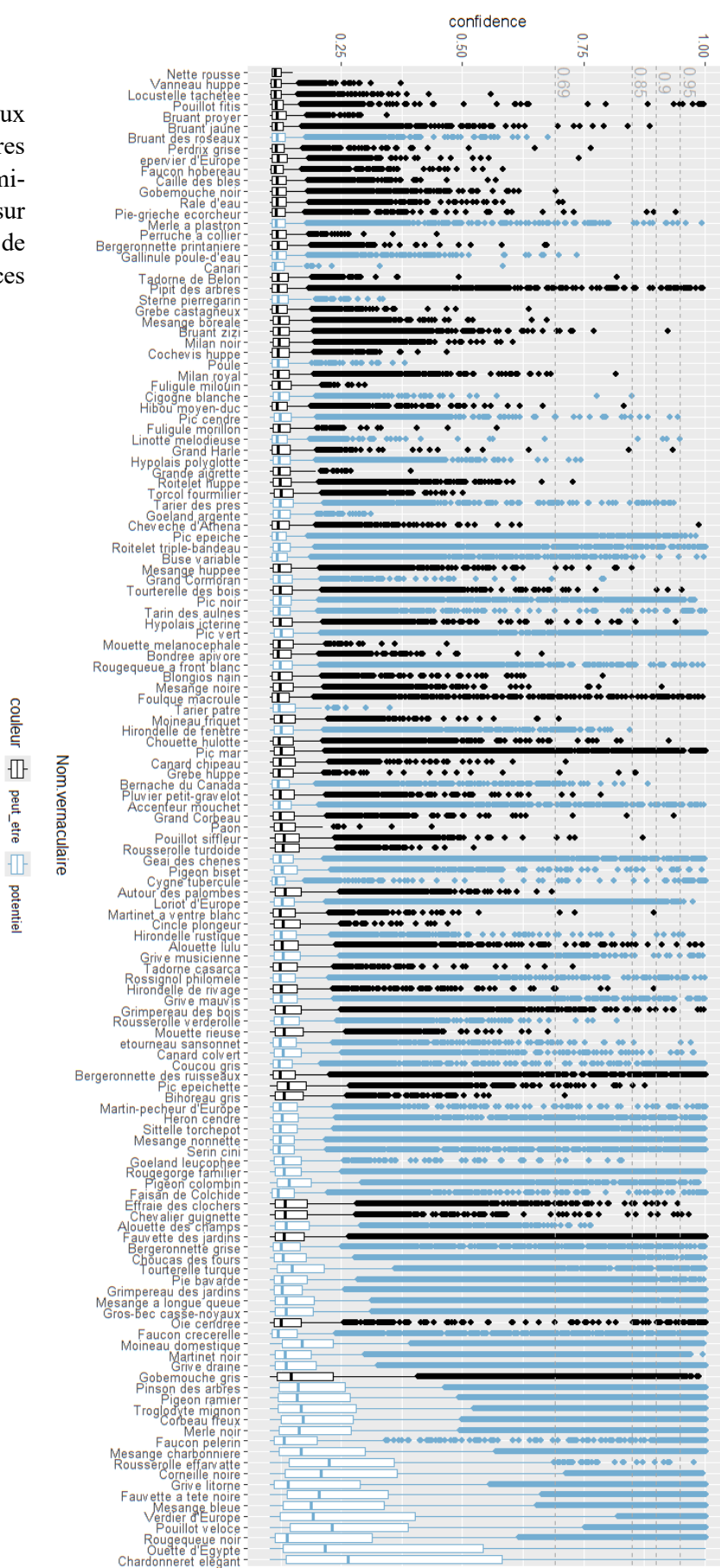
ANNEXE 1 : Liste des espèces fournie à l'algorithme au format BirdNET.

Accipiter gentilis_Northern Goshawk	Linaria cannabina_Eurasian Linnet
Accipiter nisus_Eurasian Sparrowhawk	Linaria cannabina_Eurasian Linnet
Acrocephalus arundinaceus_Great Reed Warbler	Locustella naevia_Common Grasshopper-Warbler
Acrocephalus palustris_Marsh Warbler	Lophophanes cristatus_Crested Tit
Acrocephalus scirpaceus_Eurasian Reed Warbler	Lullula arborea_Wood Lark
Actitis hypoleucos_Common Sandpiper	Luscinia megarhynchos_Common Nightingale
Aegithalos caudatus_Long-tailed Tit	Mareca strepera_Gadwall
Alauda arvensis_Eurasian Skylark	Mergus merganser_Common Merganser
Alcedo atthis_Common Kingfisher	Milvus migrans_Black Kite
Alopochen aegyptiaca_Egyptian Goose	Milvus milvus_Red Kite
Anas platyrhynchos_Mallard	Motacilla alba_White Wagtail
Anser anser_Graylag Goose	Motacilla cinerea_Gray Wagtail
Anthus trivialis_Tree Pipit	Motacilla flava_Western Yellow Wagtail
Apus apus_Common Swift	Muscicapa striata_Spotted Flycatcher
Apus melba_Alpine Swift	Netta rufina_Red-crested Pochard
Ardea alba_Great Egret	Nycticorax nycticorax_Black-crowned Night-Heron
Ardea cinerea_Gray Heron	Oriolus oriolus_Eurasian Golden Oriole
Asio otus_Long-eared Owl	Parus major_Great Tit
Athene noctua_Little Owl	Passer domesticus_House Sparrow
Aythya ferina_Common Pochard	Passer montanus_Eurasian Tree Sparrow
Aythya fuligula_Tufted Duck	Pavo cristatus_Indian Peafowl
Branta canadensis_Canada Goose	Perdix perdix_Gray Partridge
Buteo buteo_Common Buzzard	Periparus ater_Coal Tit
Carduelis carduelis_European Goldfinch	Pernis apivorus_European Honey-buzzard
Certhia brachydactyla_Short-toed Treecreeper	Phalacrocorax carbo_Great Cormorant
Certhia familiaris_Eurasian Treecreeper	Phasianus colchicus_Ring-necked Pheasant
Charadrius dubius_Little Ringed Plover	Phoenicurus ochruros_Black Redstart
Chloris chloris_European Greenfinch	Phoenicurus phoenicurus_Common Redstart
Chroicocephalus ridibundus_Black-headed Gull	Phylloscopus collybita_Common Chiffchaff
Ciconia ciconia_White Stork	Phylloscopus sibilatrix_Wood Warbler
Cinclus cinclus_White-throated Dipper	Phylloscopus trochilus_Willow Warbler
Coccothraustes coccothraustes_Hawfinch	Pica pica_Eurasian Magpie
Columba livia_Rock Pigeon	Picus canus_Gray-headed Woodpecker
Columba oenas_Stock Dove	Picus viridis_Eurasian Green Woodpecker
Columba palumbus_Common Wood-Pigeon	Podiceps cristatus_Great Crested Grebe
Corvus corax_Common Raven	Poecile montanus_Willow Tit
Corvus corone_Carrion Crow	Poecile palustris_Marsh Tit
Corvus frugilegus_Rook	Prunella modularis_Dunnock
Corvus monedula_Eurasian Jackdaw	Psittacula krameri_Rose-ringed Parakeet
Coturnix coturnix_Common Quail	Rallus aquaticus_Water Rail
Cuculus canorus_Common Cuckoo	Regulus ignicapilla_Common Firecrest
Cyanistes caeruleus_Eurasian Blue Tit	Regulus regulus_Goldcrest
Cygnus olor_Mute Swan	Riparia riparia_Bank Swallow
Delichon urbicum_Common House-Martin	Saxicola rubetra_Whinchat
Dendrocopos major_Great Spotted Woodpecker	Saxicola rubicola_European Stonechat

Dendrocoptes medius_Middle Spotted Woodpecker	Saxicola torquatus_African Stonechat
Dryobates minor_Lesser Spotted Woodpecker	Serinus canaria_Island Canary
Dryocopus martius_Black Woodpecker	Serinus serinus_European Serin
Emberiza calandra_Corn Bunting	Sitta europaea_Eurasian Nuthatch
Emberiza cirrus_Cirl Bunting	Spinus spinus_Eurasian Siskin
Emberiza citrinella_Yellowhammer	Sterna hirundo_Common Tern
Emberiza schoeniclus_Reed Bunting	Streptopelia decaocto_Eurasian Collared-Dove
Erithacus rubecula_European Robin	Streptopelia turtur_European Turtle-Dove
Falco peregrinus_Peregrine Falcon	Strix aluco_Tawny Owl
Falco subbuteo_Eurasian Hobby	Sturnus vulgaris_European Starling
Falco tinnunculus_Eurasian Kestrel	Sylvia atricapilla_Eurasian Blackcap
Ficedula hypoleuca_European Pied Flycatcher	Sylvia borin_Garden Warbler
Fringilla coelebs_Common Chaffinch	Sylvia communis_Greater Whitethroat
Fulica atra_Eurasian Coot	Sylvia curruca_Lesser Whitethroat
Galerida cristata_Crested Lark	Tachybaptus ruficollis_Little Grebe
Gallinula chloropus_Eurasian Moorhen	Tadorna ferruginea_Ruddy Shelduck
Gallus gallus_Red Junglefowl	Tadorna tadorna_Common Shelduck
Garrulus glandarius_Eurasian Jay	Troglodytes troglodytes_Eurasian Wren
Hippolais icterina_Icterine Warbler	Turdus iliacus_Redwing
Hippolais polyglotta_Melodious Warbler	Turdus merula_Eurasian Blackbird
Hirundo rustica_Barn Swallow	Turdus philomelos_Song Thrush
Ichthyaelus melanocephalus_Mediterranean Gull	Turdus pilaris_Fieldfare
Ixobrychus minutus_Little Bittern	Turdus torquatus_Ring Ouzel
Jynx torquilla_Eurasian Wryneck	Turdus viscivorus_Mistle Thrush
Lanius collurio_Red-backed Shrike	Tyto alba_Barn Owl
Larus argentatus_Herring Gull	Vanellus vanellus_Northern Lapwing
Larus michahellis_Yellow-legged Gull	

ANNEXE 2 : Niveaux de confiance associés aux détections de BirdNET pour les matinées d'enregistrement.

Les espèces bleues correspondent aux espèces observées grâce à d'autres méthodes (point d'écoute, méthode semi-automatisée) et qui sont donc présentes sur les sites de manière certaine. Il n'y a pas de certitudes quant à la présence des espèces noires.





AudioMoth dans une capsule étanche. © Héroïse Jeux, 2024

RESUME :

Ce stage s'intéresse à l'échantillonnage acoustique des communautés d'oiseaux à l'aide de l'algorithme BirdNET et des enregistreurs AudioMoth. Réalisé dans le cadre du projet Evolville à Strasbourg, l'objectif était d'évaluer la performance de BirdNET à dresser une image correcte et complète des communautés présentes le long d'un gradient urbain-forestier soumis à des paysages sonores différents. Les résultats montrent que l'allongement de la durée d'enregistrement améliore la performance de BirdNET si les données sont correctement filtrées grâce au choix d'un niveau de confiance sévère. D'autre part, l'étude a montré que les caractéristiques des sites urbains et forestiers sont essentielles à la compréhension de la performance de l'algorithme dans les paysages sonores associés, ainsi que les adaptations des espèces s'y trouvant. Cette étude est optimiste sur l'emploi de BirdNET dans des paysages sonores urbains, bien qu'il soit nécessaire d'approfondir les connaissances dans ce domaine.

Mots clés : AudioMoth, bioacoustique, BirdNET, écologie urbaine, reconnaissance automatisée d'espèces.