



HAL
open science

Protocole d'évaluation des performances des systèmes de détection-réaction. Protocole “ fournisseur ”

Alexandre Corbeau, Cyrielle Ballester, Olivier Duriez, Aurélien Besnard,
Sophie M. Dupont, Ana Maria Tobon Monroy

► To cite this version:

Alexandre Corbeau, Cyrielle Ballester, Olivier Duriez, Aurélien Besnard, Sophie M. Dupont, et al..
Protocole d'évaluation des performances des systèmes de détection-réaction. Protocole “ fournisseur ”.
Version 1.0, MSH SUD; Centre d'écologie fonctionnelle évolutive. 2023, pp.41. hal-04629050

HAL Id: hal-04629050

<https://cnrs.hal.science/hal-04629050>

Submitted on 28 Jun 2024

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

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



Distributed under a Creative Commons Attribution 4.0 International License

Protocole d'évaluation des performances des systèmes de détection-réaction ;

Protocole « fournisseur »

Dupont Sophie M., Corbeau Alexandre, Ballester Cyrielle,
Duriez Olivier & Besnard Aurélien

Centre d'Écologie Fonctionnelle et Evolutive, Univ Montpellier, CNRS UMR5175, EPHE-PSL University, IRD, Montpellier, France – équipe de recherche du projet MAPE (Mortalité Aviaire dans les Parcs Éoliens terrestres en exploitation)

Citation préconisée : Dupont S. M., Corbeau A., Ballester C., Duriez O. & Besnard A. (2023) – Protocole d'évaluation des systèmes de détection-réaction, protocole « fournisseur ». 41 pages.

Version 1.0, dernière mise à jour le 10/05/2023.

Table des matières

TABLE DES MATIERES	2
ABREVIATIONS	3
INTRODUCTION	4
1. LE PARAMETRE A ESTIMER ET LES VARIABLES D'INFLUENCE	6
1.1 Probabilité de détection/classification	6
1.2 Le choix des moyens-tiers pour estimer ce paramètre	8
1.3 Les variables d'influence	13
• L'espèce détectée	13
• La distance de détection/classification	14
• L'arrière-plan lors de la détection/classification	15
• Conditions météorologiques	16
• Eclairage lumineux lors de la détection/classification	17
• L'angle de vue par le système de détection-réaction	17
• L'azimut de détection	18
2. ACQUISITION DES DONNEES SUR LE TERRAIN	18
2.1 Choix d'un lieu et d'une période	18
2.2 L'installation des systèmes de détection-réaction	20
2.3 Mesure des variables d'influence environnementales	22
2.3.1. Définition des conditions météorologiques et de l'éclairage lumineux	22
2.3.2. Définition de l'arrière-plan	23
2.4 Acquisition des données des cibles via un moyen-tiers	23
2.4.1. Utilisation d'individus sauvages équipés de balises GPS	23
2.4.2. Utilisation d'observateurs humains équipée d'un télémètre laser	24
3. TRAITEMENT ET MODELISATION STATISTIQUE DES DONNEES ACQUISES SUR LE TERRAIN	29
4. TAILLES D'ECHANTILLON ET PRECISION DES ESTIMATIONS	32
RÉFÉRENCES	35
ANNEXE 1 - MODELE DE GRILLE FINALE DES PERFORMANCES ESTIMEES POUR CHAQUE SDA PAR APPLICATION DU PROTOCOLE « FOURNISSEUR »	36
ANNEXE 2 - INTERVALLES DE CONFIANCE A 95% OBTENUS EN FONCTION DU NOMBRE DE TRAJECTOIRES UTILISE ET DE LA PROBABILITE DE DETECTION DU SDA	39

Abréviations

. Degré Celsius : °C

. Degré d'angle : °

. Heure : h

. Kilomètre : km

. Litre : L

. Lux : lx

. Mètre : m

. Mil angulaire : mil

. Millimètre : mm

. Minute : min

SDA : Système de détection-réaction

Seconde : s

. Watt : W

. Work Package : WP

Introduction

Ce document a été rédigé dans le cadre du programme de recherche collaboratif MAPE (<https://mape.cnrs.fr/fr>, « Réduction des Mortalités Aviaires dans les Parcs Éoliens terrestres en Exploitation »). Il résulte du travail réalisé dans le cadre du WP4 dont l'objectif était de rédiger un protocole d'évaluation des performances des systèmes de détection-réaction installés dans les parcs éoliens terrestres pour réduire les collisions d'oiseaux.

Dans ce cadre, deux protocoles ont été élaborés afin d'évaluer les **performances** des systèmes de détection-réaction (désignés par l'abréviation **SDA** dans la suite du document). Ces protocoles n'ont pas vocation à évaluer l'**efficacité** (définie comme la capacité de réduction des mortalités) des SDA (voir la justification de ce choix dans Corbeau et Besnard, 2021). Ils n'ont pas non plus vocation à identifier les sources de pannes ou les divers problèmes de réglage des SDA. Ce choix de la question scientifique à traiter dans les protocoles MAPE a été fait au cours des premiers séminaires dédiés au projet, séminaires rassemblant l'ensemble des parties prenantes impliquées dans la conciliation du développement de l'éolien en France et de la préservation de la Biodiversité. Ces protocoles ont par la suite été co-construits avec l'ensemble des membres du comité technique du Work Package 4 (WP4) de ce programme de recherche. Ce comité technique inclut, là aussi, des membres de toutes les parties prenantes impliquées dans les questions du développement de l'éolien en France et de la préservation de la biodiversité : les opérateurs de parcs éoliens, les bureaux d'études, les associations de protection de la nature, les DREALs, l'OFB et des institutions de recherche. Les fournisseurs de SDA ont volontairement été écartés de ce comité technique pour ne pas générer de conflit d'intérêt. Cependant, ils ont tous été consultés afin d'obtenir des informations précises sur le fonctionnement et les caractéristiques technologiques des dispositifs qu'ils déploient actuellement sur le territoire français.

Ce document fait suite à la rédaction, dans le cadre de la première année de MAPE, d'une note de synthèse concernant l'évaluation des SDA, diffusée en août 2021 (Corbeau et Besnard, 2021) intitulée « *Les systèmes de détection-réaction dans les parcs éoliens, un moyen de réduction des mortalités aviaires : Principes et conseils pour une bonne évaluation* ». Cette note est téléchargeable sur le site internet du programme MAPE (<https://mape.cnrs.fr/wp-content/uploads/2021/12/Corbeau-Besnard-MAPE-2021-Principes-et-conseils-pour-une-bonne-evaluation-des-sytemes-de-detection-reaction.pdf>). Les protocoles rédigés par la suite dans le cadre du WP4 de MAPE découlent de fait directement des recommandations émises dans cette note de synthèse.

L'objectif des deux protocoles rédigés dans le cadre de MAPE est d'obtenir une évaluation standardisée, homogène et neutre des performances des SDA déployés sur les parcs pour réduire les mortalités d'oiseaux par collision. Ils se veulent applicables à tous les dispositifs actuellement déployés sur le territoire français (à la date de publication de ce document, deuxième trimestre 2023), indépendamment de la technologie employée, mais aussi aux systèmes qui seront développés et mis sur le marché dans le futur. A noter que ces protocoles visent à une évaluation diurne des systèmes. En effet, à l'heure actuelle, il n'existe pas d'approche expérimentale robuste pour mesurer les performances nocturnes de ces dispositifs, notamment du fait de l'absence de moyen-tiers pertinent la nuit (voir partie 1.2. pour le choix des moyens-tiers retenus dans le cadre des protocoles développés par l'équipe scientifique de MAPE).

L'évaluation des performances des SDA peut avoir deux objectifs sensiblement différents (i) porter à connaissance ces performances génériques pour que les opérateurs éoliens puissent les comparer et choisir le dispositif le plus adapté à leur parc, (ii) évaluer si ces performances respectent les obligations réglementaires en vigueur dans un parc donné. De ce fait, il a été décidé d'élaborer deux protocoles distincts et complémentaires : un protocole « fournisseur » à destination des entreprises fournisseurs de systèmes, qui répond à l'objectif de porter à connaissance les performances génériques des systèmes en testant leurs performances dans des situations variées ; et le protocole « parc » à destination des exploitants de parcs éoliens permettant de tester un SDA sur un parc donné afin de vérifier que ses performances correspondent bien aux obligations réglementaires. Dans le présent document, seul le protocole « fournisseur » est présenté. Le protocole « parc » fait l'objet d'un document séparé, disponible, lui aussi, sur le site web de MAPE.

Pour que la démarche soit pertinente pour les opérateurs qui veulent s'équiper d'un système, il est important qu'ils puissent baser leur choix de SDA sur une comparaison des performances des différents systèmes disponibles sur le marché. Ainsi, le protocole « fournisseur » doit être mis en œuvre de manière identique et systématique pour chaque modèle de SDA commercialisé par une entreprise « fournisseur ». Contrairement au protocole « parc » qui s'intéresse aux quatre phases d'opérationnalité des SDA (le fonctionnement du système, la détection des cibles, leur classification et la réaction du système et de l'éolienne, Corbeau et Besnard [2021]), le protocole « fournisseur » se concentre sur les étapes de détection et de classification. Ce sont en effet les deux seules étapes qui sont indépendantes du parc éolien d'implantation (paysage environnant et caractéristiques techniques des éoliennes). Elles seront évaluées de manière approfondie pour chaque système, dans un même lieu, éventuellement sans éolienne.

L'évaluation via ce protocole « fournisseur » a pour but d'établir une grille des performances de détection et de classification des différents SDA, et cela pour chaque modalité des facteurs influençant potentiellement les performances de ces deux étapes (Annexe 1). Dans l'idéal, toutes les modalités de ces facteurs devraient être testées. Néanmoins, les fournisseurs peuvent décider de ne pas en tester certaines, selon les capacités *a priori* de leurs systèmes. Par exemple, certains systèmes ne sont pas configurés pour

détecter des objets volants à une distance excédant 600 m. De fait, le fournisseur peut décider de ne pas réaliser de tests au-delà de cette portée de détection. Il sera alors précisé dans la grille de performances que ceci n'a pas été évalué (au moyen de la valeur « NA » par exemple, pour « Not Available »). Les grilles finales de performances ainsi obtenues pourront être mises à disposition des opérateurs pour les aider à choisir le système qui leur semble le plus adapté à leur contexte (espèce cible, paysage, budget, etc.).

1. Le paramètre à estimer et les variables d'influence

1.1 Probabilité de détection/classification

Schématiquement, les risques de collision surviennent lorsque des oiseaux passent dans l'environnement proche des éoliennes. Les oiseaux se déplacent en trois dimensions (longitude, latitude, altitude) à proximité des turbines et leurs trajectoires peuvent, à un moment donné, intersecter la zone de rotation des pâles. Les SDA ont alors pour objectif de détecter les individus en vol ; de classer, à la fois l'oiseau (souvent selon sa taille sans nécessairement classer la cible à l'espèce) et sa trajectoire, comme étant « à risque » ou pas et de mettre en place une réaction appropriée (effarouchement sonore ou visuel ou arrêt de la turbine), le cas échéant. Pour évaluer les performances des SDA, il convient, de fait, d'estimer les probabilités de détection et de classification des espèces concernées (ou des espèces de taille similaire) par les arrêtés préfectoraux en vigueur dans les parcs français.

Il semble cependant difficile d'estimer séparément les probabilités de détection et de classification pour tous les systèmes déployés en France aujourd'hui. En effet, la majorité des systèmes actuellement développés recherche constamment les objets volants ou les variations de pixel dans les images acquises (étape de détection d'objets potentiellement pertinents). Ces systèmes détectent en continu des objets, qui sont « classés » en continu également. Cette évaluation de la nature des cibles détectées (étape de classification) se réalise de façon quasi-instantanée (avec des vitesses de calcul souvent très importantes), afin de les exclure immédiatement (dans la majorité des cas) ou au contraire de poursuivre vers une qualification plus précise du risque. Dans ce contexte, il paraît donc complexe de séparer les processus de détection et de classification. Par ailleurs, les données sauvegardées par les SDA ne concernent, en général, que les événements détectés et classifiés comme pertinents (ayant induit une réaction ou non, *informations fournies par les fournisseurs*). Ceci rend, de fait, impossible la séparation de ces deux étapes à partir des données enregistrées par défaut. Les fournisseurs nous ont bien précisé qu'il est, en théorie, possible de conserver des informations pour tous les événements détectés (classifiés ou non ; ayant induit une réaction

ou non). Néanmoins, leur extraction représente un travail conséquent et nécessite une capacité de stockage très importante (certains fournisseurs ont comptabilisé plusieurs dizaines de milliers d'événements détectés sur une journée). Le traitement de telles données implique aussi un temps d'analyse considérable puisqu'il faudrait visionner « manuellement » l'ensemble des vidéos de détections pour déterminer si la classification a été pertinente. Par ailleurs, les fournisseurs considèrent que l'évaluation séparée des probabilités de détection et de classification n'apporte pas d'élément particulièrement pertinent à l'utilisateur potentiel de ces systèmes puisque c'est bien le résultat des deux étapes qu'il est nécessaire d'évaluer pour savoir si le système prend les bonnes décisions de déclenchement d'une réaction. Du fait de ces éléments, **le protocole « fournisseur » se focalise sur la probabilité de détection/classification combinée des SDA.**

A partir de la trajectoire d'un oiseau, plusieurs variables d'intérêt associées aux performances des systèmes ont été utilisées dans la littérature scientifique ou dans les retours d'expérience antérieurs de tests de performance des SDA. L'une des plus couramment employées est la distance moyenne oiseau-système de première détection/classification par le dispositif. Cet estimateur atteste des capacités du système à détecter/classifier un oiseau à une certaine distance (voir par exemple McClure et al., 2018). Cependant, ce paramètre ne permet pas de connaître la proportion de trajets « à risque » que le système détecte/classifie comme tel. Or, certains arrêtés préfectoraux demandent explicitement aux opérateurs de parcs éoliens de démontrer que XX% des trajets sont détectés/classifiés avant une certaine distance seuil. De plus, c'est bien cette capacité à détecter/classer avant une certaine distance considérée comme à risque qui atteste que le système pourra permettre d'éviter des collisions (voir le WP3-R3 du programme MAPE pour des réflexions sur ces distances et l'application [eoldist](#) si ces distances ne sont pas précisées dans l'arrêté).

Ainsi, nous avons choisi, pour ce protocole, **la probabilité de détection/classification combinée que le système détecte un oiseau avant une certaine distance** comme paramètre à estimer pour quantifier les performances des SDA. Ce paramètre semble plus informatif et plus en adéquation avec les demandes réglementaires énoncées dans les arrêtés préfectoraux que les autres paramètres décrits dans la littérature. En effet, c'est la valeur de référence pour déterminer si le système détecte et classe bien les objets à risque à temps. Cette probabilité de détection/classification est estimée entre 0 et 1 000 m autour de l'éolienne en utilisant tous les trajets enregistrés par le moyen-tiers (voir partie suivante 1.2. sur le choix du moyen-tiers) pour chaque distance seuil (figure 1).

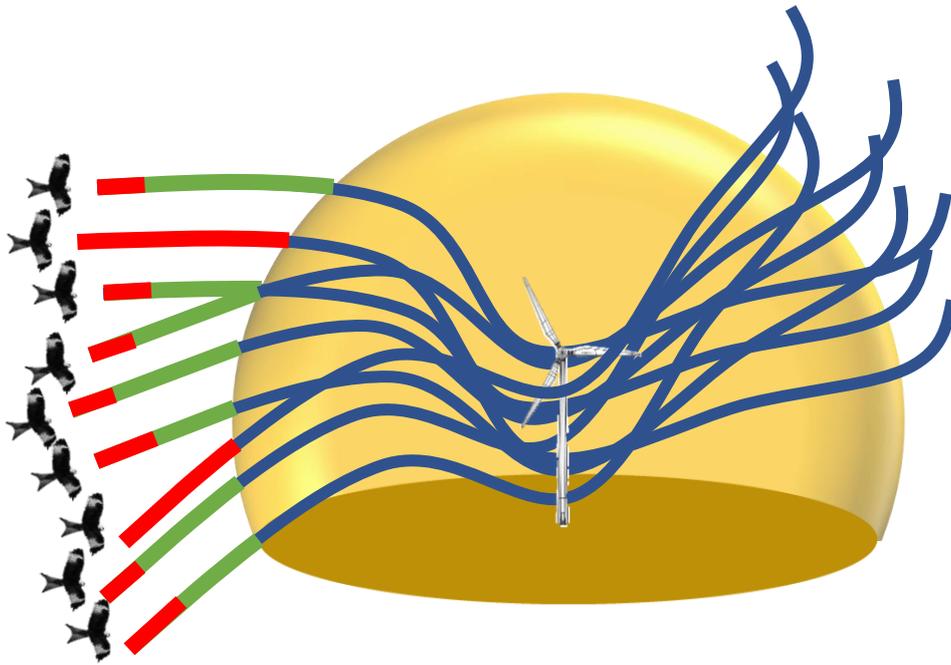


Figure 1 : Estimation de la probabilité de détection/classification combinée avant une certaine distance (sphère jaune) à partir des trajectoires relevées sur le terrain. La trajectoire prend la valeur « 1 » si elle est détectée/classifiée avant cette distance (portion de trajectoire verte) et « 0 » si elle n'est pas détectée/classifiée (portion de trajectoire rouge) avant cette distance.

1.2 Le choix des moyens-tiers pour estimer ce paramètre

L'estimation de la probabilité de détection/classification combinée avant une certaine distance implique de disposer :

- 1) Des données relatives aux évènements qui ont été détectés et classifiés par les SDA.
- 2) Des géolocalisations d'oiseaux relevées par un moyen-tiers fiable et robuste.

La difficulté réside dans le fait de disposer de moyens-tiers qui soient les plus complets et les plus exhaustifs possibles pour servir de référence (voir la note de synthèse de Corbeau et Besnard [2021] qui illustre le biais d'évaluation lorsque ces critères ne sont pas remplis). Dans le cadre de ce protocole, deux moyens-tiers ont été retenus : le relevé de trajectoires d'individus sauvages d'espèces cibles équipés de balises GPS et le relevé de trajectoires d'oiseaux sauvages survolant le parc éolien d'intérêt par une paire d'observateurs humains équipée d'un télémètre laser.

Dans la note de synthèse de Corbeau et Besnard (2021), deux autres moyens-tiers ont également été évoqués : les drones et les oiseaux de fauconnerie. Néanmoins, la poursuite de nos échanges avec les fournisseurs de SDA, des télépilotes et des fauconniers professionnels, ainsi que la rencontre de chercheurs

internationaux et chargés de mission de bureaux d'étude appliquant sur le terrain des protocoles d'évaluation des performances de ces dispositifs de réduction des mortalités aviaires ont fait évoluer notre réflexion concernant ces deux moyens-tiers.

- **L'utilisation de drones**

De prime abord, l'utilisation de drones avec géolocalisation 3D embarquée peut être considérée comme une approche précise des performances des SDA. Elle permet en effet d'acquérir une grande quantité de données sur un temps limité et assure un contrôle quasi-total de la trajectoire de la cible volante testée. Ceci assure d'avoir des estimations précises du fait de grandes tailles d'échantillons. Cependant, les drones ne miment pas les oiseaux. Ils n'ont pas la même taille, pas la même forme, pas la même réflectance, pas le même type de trajet (etc.) que des oiseaux. Il n'y a de ce fait aucune garantie que les algorithmes de classification des systèmes aient les mêmes taux de détection/classification pour des drones que pour des oiseaux. Par ailleurs, plus les algorithmes de classification évolueront, notamment via l'utilisation d'intelligence artificielle, plus ils auront tendance à exclure les drones comme des objets pertinents, ce qui conduira à biaiser négativement les performances des systèmes. A noter que certains algorithmes ont déjà été entraînés pour reconnaître les drones. Cependant, même dans cette situation, rien ne permet de certifier que la probabilité de détection/classification est identique pour les drones et les espèces cibles. L'ensemble de ces éléments fait qu'il est probable que les probabilités de détection/classification obtenues avec des drones soient biaisées par rapport à celles attendues avec des oiseaux sauvages. Cette approche a donc été exclue du protocole.

- **L'utilisation d'oiseaux de fauconnerie**

Les oiseaux de fauconnerie permettent d'appréhender de façon contrôlée et donc exhaustive la présence de cibles volantes dans l'environnement du SDA. Des vols peuvent être organisés pour acquérir des géolocalisations 3D sur la majeure partie du gradient de chaque variable d'influence (i.e. distance de détection/classification, arrière-plan, visibilité, pluviométrie, nébulosité, luminosité, angle d'incidence du soleil, azimut, voir partie 1.3.). Il est possible de faire voler ces oiseaux dans des conditions météorologiques variées (entre 0°C et 40°C ; avec une pluie modérée ; avec un brouillard peu dense ; sous des vitesses de vent allant jusqu'à 27,8 m/s). De plus, l'équipement de ces oiseaux de balises de géolocalisation, relevant toutes les secondes la position 3D de l'oiseau, permet une analyse fine et poussée des performances de détection/classification des dispositifs. Cette méthode présente également l'avantage de mobiliser de vrais oiseaux. Il est par ailleurs possible de disposer d'oiseaux d'espèces de tailles variées et ciblées par des arrêtés préfectoraux dans certains parcs éoliens. La classification par les algorithmes est alors probablement représentative des oiseaux sauvages.

Toutefois, des limites rédhibitoires à l'utilisation de ce moyen-tiers pour les tests de performance ont été identifiées suite à la consultation de fauconniers professionnels. En effet, même en réalisant des tests avec des espèces dites de « haut vol » (faucon pèlerin, aigle royal ; à différencier des espèces de « bas-vol » comme les autours, faucon sacre, buse de Harris), les oiseaux ne sont généralement pas entraînés pour se déplacer à des hauteurs supérieures à 10 m. Par ailleurs, ces oiseaux de chasse sont généralement entraînés pour voler en ligne droite, à vitesse constante (~ entre 60 et 90 km/h, selon les espèces) et sur quelques dizaines de mètres maximum. Du fait de ces contraintes, il est rarement possible de générer des trajectoires variées et imprévisibles, proches de celles des oiseaux sauvages. Il n'est également pas possible de réaliser des vols au crépuscule et à l'aube car si un oiseau s'éloigne du site de lâcher (qui lui est inconnu) durant ces phases journalières de faible luminosité/visibilité, il y a un risque qu'il ne parvienne pas à revenir auprès de son fauconnier. Pourtant, ce sont des périodes à fort intérêt dans l'évaluation des performances des systèmes de détection-réaction justement du fait de la faible luminosité présente (voir partie 1.3.). De fait, le comportement de vol de ces oiseaux est potentiellement assez éloigné de celui d'oiseaux sauvages, ce qui peut biaiser leur détectabilité/classification par le dispositif de réduction des mortalités aviaires. Du fait de ces différentes contraintes cette approche a, elle aussi, été exclue du protocole.

- **L'utilisation d'individus sauvages équipés de balises GPS**

Certains individus d'espèces d'intérêt dans le cadre de la problématique éolienne utilisent la zone de test du SDA comme zone de chasse ou de reproduction et peuvent alors être directement utilisés comme moyen-tiers de référence, en équipant ces individus de balises GPS à haute résolution. De fait, ils survoleront régulièrement et naturellement la zone d'intérêt et permettront alors d'acquérir passivement une grande quantité de données de géolocalisation, sur une échelle temporelle très fine (longitude, latitude, altitude toutes les secondes). Cependant, il faut noter que la pose de telles balises géolocalisatrices à résolution temporelle fine n'est possible actuellement que sur des espèces de grandes et moyennes tailles tels que les aigles, les vautours ou les milans du fait du poids de ces appareils. Cette approche ne permettra donc pas de compléter la grille de performance pour de plus petites espèces (à noter qu'avec les progrès technologiques, il n'est pas exclu d'équiper des petits faucons avec des balises GPS similaires dans les années à venir). A noter aussi que déployer ce type de technologie implique, en particulier lorsqu'elle concerne plusieurs espèces différentes sur un même site, des autorisations parfois complexes à obtenir (Permis de capture d'oiseau validé annuellement par le Centre de Recherche sur la Biologie des Populations d'Oiseaux, au Museum National d'Histoire Naturelle de Paris). Par ailleurs, cette approche méthodologique est relativement coûteuse (l'étape de capture d'un individu donné pour son équipement implique de mobiliser des spécialistes et est chronophage). Cette approche est donc très pertinente au vu de la quantité/qualité des données qu'elle peut fournir mais présente des limites certaines qui nécessitent une

évaluation approfondie. Dans l'éventualité où un site expérimental national serait déployé, il pourrait alors sûrement être envisagé d'équiper une gamme d'espèces variées pour la tester.

- **L'utilisation d'une paire d'observateurs équipée d'un télémètre laser**

La mobilisation simultanée de deux observateurs humains complémentaires sur le terrain et l'utilisation de jumelles à télémètre laser intégré sont les deux conditions *sine qua non* pour que ce moyen-tiers puisse être considéré comme robuste et fiable. En effet, comme souligné dans Corbeau et Besnard (2021), individuellement, les humains possèdent des capacités de détection généralement plus faibles que celles des SDA (McClure et al., 2018). De ce fait, ni le SDA, ni les observateurs humains ne sont exhaustifs. Ceci pose problème pour estimer les probabilités de détection/classification combinées du SDA avant une certaine distance. En effet, certains trajets à risque pourront avoir été ratés par les deux types de systèmes d'observation.

Cette difficulté de détection non-exhaustive des deux systèmes d'observation est un problème classique en recherche écologique qui a été exploré dans les années 1960 pour estimer les effectifs d'une population via deux comptages indépendants. La méthode utilisée dans ce contexte s'appelle l'indice de « Lincoln-Peterson ». Elle implique que les individus soient identifiés, c'est-à-dire qu'il est possible de déterminer quels sont ceux qui n'ont été détectés que par un seul système d'observation et ceux qui au contraire ont été détectés par les deux systèmes d'observation simultanément. Le raisonnement à la base de cette méthode peut être adapté au problème de l'estimation de la probabilité de détection/classification des SDA avec ici comme individus les trajets d'oiseaux transitant par la zone d'intérêt (box 1). Pour être pertinente, cette approche implique que la probabilité de détection des observateurs soit indépendante de la probabilité de détection par le SDA.

Etant donné que tous les SDA ne mesurent pas la distance entre l'objet détecté et le dispositif, il est aussi obligatoire que le moyen-tiers de référence palie à ce manque en acquérant cette donnée de manière précise. Les jumelles à télémètre laser intégré sont un appareil de mesure permettant d'obtenir d'une part la distance entre l'objet visé et le télémètre laser mais aussi de collecter des informations sur la géolocalisation de l'objet ciblé ainsi que l'azimut de mesure. Ce relevé de données par les observateurs humains équipés d'un télémètre laser peut donc suffire pour estimer la probabilité de détection/classification combinée avant une certaine distance même en absence de mesure de distance par le SDA.

Cette approche via des observateurs humains présente l'avantage d'être déclinable dans n'importe quel site et pour n'importe quelle espèce.

Box 1 – Application de l'indice de Lincoln-Peterson pour estimer la probabilité de détection/classification combinée avant une certaine distance

La méthode utilisée classiquement en écologie pour estimer le nombre total d'objets présents lorsque **deux observateurs indépendants** ne sont pas exhaustifs s'appelle l'indice de « Lincoln-Peterson ». Elle implique que les « individus » (dans le cas qui nous intéresse ici, les trajets suivis) soient identifiés, c'est-à-dire qu'il est possible de déterminer quels sont ceux qui n'ont été détectés que par un seul système d'observation et ceux qui, au contraire, ont été détectés par les deux systèmes d'observation simultanément.

L'indice de Lincoln-Peterson se formule ainsi : $N_{tot} = \frac{(N_{obs}+1) \times (N_{sys}+1)}{N_{com}+1}$ (eq. 1)

Avec N_{obs} = nombre de trajets détectés par l'observateur humain avant une certaine distance (connu car mesuré par l'observateur), N_{sys} = nombre de trajets détectés par le système avant une certaine distance (pas systématiquement connu car certains systèmes ne mesurent pas les distances aux cibles détectées), N_{tot} = nombre total de trajets avant cette distance (pas connu car certains trajets peuvent être ratés par les deux systèmes d'observation), N_{com} = nombre de trajets détectés en commun (observateur humain + système, connu grâce à la mise en correspondance des données de détection/classification relevées par le SDA avec celles obtenues par les observateurs humains).

Le principe de cet indice repose sur l'idée que la proportion de trajets détectés par le SDA au sein des trajets détectés par les observateurs humains équipés d'un télémètre laser est une estimation non-biaisée de la proportion de trajets détectés par le SDA au sein de l'ensemble des trajets qui ont eu lieu lors du test. L'estimation du nombre total de trajets s'obtient donc par une simple règle de trois.

Dans le cas des SDA, nous ne cherchons pas particulièrement à avoir une estimation du nombre total de trajets mais simplement de la probabilité de détection/classification combinée par le système avant une certaine distance. Le principe est donc encore plus simple. Ainsi, si on pose : P_{obs} = Probabilité de détection/classification combinée d'un oiseau par l'observateur humain avant une certaine distance, P_{sys} = même chose pour le SDA (paramètre que l'on cherche à estimer),

On peut donc écrire : $P_{obs} = \frac{N_{obs}}{N_{tot}}$ (eq. 2) ; $P_{sys} = \frac{N_{sys}}{N_{tot}}$ (eq. 3) ; $N_{com} = N_{tot} \times P_{obs} \times P_{sys}$ (eq. 4)

En réorganisant les eq. 3 et eq. 1, on obtient $N_{com} = N_{obs} \times P_{sys}$ (eq. 5)

Soit tout simplement $P_{sys} = \frac{N_{com}}{N_{obs}}$ (eq. 6)

La variance de cet estimateur peut être facilement obtenue par une modélisation de type modèle linéaire généralisé avec une distribution binomiale des données.

Il est important de savoir que la précision de l'estimation dépend fortement des probabilités de détection des deux systèmes : plus ces probabilités sont élevées, plus les estimations sont précises.

L'hypothèse que les deux systèmes d'observation sont indépendants est cruciale pour pouvoir écrire l'équation 4 (multiplication des probabilités de détection des deux systèmes d'observation).

1.3 Les variables d'influence

Les variables décrites ci-après ont été retenues car elles peuvent impacter les probabilités de détection/classification combinées avant une certaine distance. On parle de « variables explicatives » en langage statistique. Elles sont, *a priori*, indépendantes entre elles (c'est-à-dire que la variation de valeur de l'une de ces variables ne modifie pas l'effet engendré par une autre), à l'exception probablement de la variable « espèce » qui doit interagir avec de nombreuses autres variables d'influence (voir ci-dessous). La grille de performance du système doit donc prendre en compte les différentes modalités de ces variables pour décrire précisément les conditions dans lesquelles les performances sont plus ou moins bonnes.

- **L'espèce détectée**

L'espèce cible joue un rôle majeur dans la capacité de détection/classification du SDA, notamment via la taille des individus. Plus précisément, un oiseau de grande envergure est sans aucun doute détectable/classifiable plus aisément (à une plus grande distance, dans des conditions de luminosité plus faibles, etc.) qu'un oiseau de petite taille. Cette variable est, de fait, abordée dans la grille d'évaluation par un découpage de la grille selon les gammes de taille présentées dans la box 2. Dans l'étape de modélisation des données, il s'agira donc d'une variable catégorielle à trois modalités : « Oiseaux de grande taille », « Oiseaux de taille moyenne », « Oiseau de petite taille » qui correspondent respectivement aux oiseaux d'envergure supérieure à 150 cm, aux oiseaux d'envergure comprise entre 100 cm et 150 cm et aux oiseaux d'envergure inférieure à 100 cm (box 2). A noter que certains systèmes classifient les objets détectés à l'espèce. Pour ces systèmes, il ne sera donc possible d'évaluer les performances que pour les espèces reconnues. Ceci devra être mentionné dans la grille finale de performances.

Box 2 – Espèces proches pouvant être regroupées

Dans le cadre du protocole « fournisseur » il est demandé d'évaluer les performances pour différentes catégories d'espèces en fonction de leur taille. Cette taille est en effet un facteur important pouvant affecter la probabilité de détection/classification. Trois classes de taille sont proposées :

- Les oiseaux de grande taille (envergure supérieure à 150 cm) tels que les vautours, l'aigle royal, le pygargue à queue blanche, les deux cigognes ou la grue cendrée, le circaète Jean-le-blanc, l'aigle de Bonelli, le balbuzard pêcheur, hérons cendrés, cygnes et oies.
- Les oiseaux de taille moyenne (envergure comprise entre 100 cm et 150 cm) tels que l'aigle botté, le milan royal, le milan noir, la buse variable, la bondrée apivore, l'autour des palombes ou les trois busards, les hérons et pourprés, le grand corbeau, les goélands.
- Les autres oiseaux de petite taille (envergure inférieure à 100 cm) tels que tous les faucons, l'élanion blanc ou l'épervier d'Europe, les autres corvidés, les pigeons.

• La distance de détection/classification

Il existe une forte corrélation négative entre la distance oiseau-SDA et la probabilité de détection/classification par le système. Concrètement, plus une cible se trouve loin du dispositif, plus sa taille sur l'image est réduite, et moins bien elle sera potentiellement détectée/classifiée par le système. Cette variable est une variable continue. Cependant, dans le cadre de ce protocole, nous avons fait le choix de la discrétiser, c'est-à-dire de la diviser en classes comprises entre 0 et 1 000 m. Ce choix repose sur le fait qu'elle sera plus facile à manipuler dans les analyses sous cette forme plutôt que sous forme continue. En effet, l'influence de la distance oiseau-SDA peut ne pas être linéaire et des effets de seuil peuvent notamment être rencontrés. Par ailleurs, l'objectif étant d'estimer la probabilité de détection/classification combinée avant une certaine distance, l'idée de discrétiser cette variable est cohérente avec sa représentation dans la grille finale de performances (Annexe 1). Enfin, des essais d'analyse de données simulées nous ont montré que les modèles pouvaient avoir du mal à estimer la pente de la relation entre distance et probabilité de détection/classification dans le cas où la variable est considérée sous forme linéaire. Ici, nous avons retenu des classes de 100 m. En effet, il paraît difficilement réaliste de descendre à des échelles plus fines sachant la résolution des temps de relevé avec un télémètre laser (un point est enregistré toutes les 5 s en moyenne, pour un observateur humain entraîné à l'utilisation de cet appareil). A l'inverse, des classes de distances plus larges auraient peu de sens en termes d'évaluation des performances. Il est important de noter que cette

variable est bien la distance entre l'oiseau et le SDA. De fait, dans le cas de l'utilisation de jumelles à télémètre laser intégré, il faudra calculer cette distance oiseau-SDA en utilisant les coordonnées 3D de l'oiseau et du SDA (voir ci-dessous la partie 3.).

- **L'arrière-plan lors de la détection/classification**

D'après les fournisseurs de SDA optiques, un ciel bleu homogène (fond uni sans mouvement intrinsèque) permet d'augmenter les contrastes entre l'arrière-plan et un objet volant. Les capacités de détection/classification du dispositif seraient alors maximales dans un tel contexte d'arrière-plan.

De la même manière, les objets se déplaçant sur un fond polychrome sans mouvement (sol sans végétation, falaise ou ciel nuageux par exemple) seraient détectés/classifiés avec une plus grande probabilité par les caméras que ceux se déplaçant devant un fond polychrome composé de végétation haute (arbres, culture, prairies, haies) présentant des mouvements. La prise en compte de ce facteur d'influence pour la modélisation de la probabilité de détection/classification combinée du dispositif avant une certaine distance semble donc essentielle.

A noter que les signaux provenant d'un arrière-plan en mouvement (notamment de la végétation au sol) peuvent aussi interférer avec les ondes reçues par les systèmes utilisant la technologie radar. De fait, cette variable est importante à intégrer dans les analyses, que cela soit pour les systèmes optiques ou les systèmes radars.

Cette variable peut être complexe à caractériser notamment car, lors de la mobilisation d'observateurs humains avec télémètre laser, l'observateur n'est pas nécessairement positionné exactement au même emplacement (localisation géographique et angle de vue) que le SDA évalué. L'arrière-plan relevé par l'observateur n'est alors pas celui observé par le SDA, ce qui peut alors altérer les données collectées et de fait les analyses associées. En particulier, pour les SDA qui sont installés sur le mât des éoliennes (en hauteur), une distance minimale entre le télémètre laser et le système est obligatoire car l'éolienne peut perturber le fonctionnement de l'appareil de mesure (voir ci-dessous, partie 2.4.2.). Lorsque ce sont des oiseaux sauvages équipés de balises GPS qui sont impliqués dans l'évaluation des performances des SDA, l'arrière-plan n'est alors tout simplement pas disponible car pas relevé par le moyen-tiers mobilisé.

Il est possible de caractériser l'arrière-plan de trajectoire d'un oiseau enregistrée par le télémètre ou un GPS grâce à une cartographie 3D de l'environnement du SDA. Cette cartographie 3D peut être obtenue en utilisant un MNT (Modèle Numérique de Terrain) à 5 m de précision. L'arrière-plan peut alors être défini en utilisant l'alignement entre le SDA et la localisation de l'oiseau détecté. Un script R a été développé par l'équipe de MAPE à cet effet (téléchargeable depuis le site de MAPE).

Concrètement, avec cette approche, les coordonnées géographiques du point d'arrière-plan sont déterminées. Il est donc en théorie possible de définir le type de végétation occupant cette coordonnée (par exemple avec une cartographie d'occupation du sol de type CORINE Land Cover) pour préciser si cet arrière-plan est susceptible de présenter des mouvements ou pas. Cependant, il est peu probable que de simples catégories d'occupation du sol suffisent pour pouvoir déterminer s'il y a présence de mouvement au pas (par exemple, les « zones agricoles » peuvent être des cultures avec mouvements mais ne plus présenter de mouvement à la suite des récoltes). C'est pourquoi nous avons fait le choix de restreindre la variable catégorielle « arrière-plan » à deux modalités, à savoir « fond de végétation » ou « fond de ciel ».

A noter que le programme national LIDAR HD en cours d'acquisition (à la date de rédaction de ce rapport) par l'IGN pourrait fournir un MNS (Modèle Numérique de Surface) plus précis qu'un MNT à 5 m. En revanche, pour le moment, seule une très faible proportion du territoire français a été cartographiée à l'aide de cet outil.

- **Conditions météorologiques**

Les mauvaises conditions météorologiques (nuage, brouillard, pluie, neige, grêle, etc.) induisent une diminution de la visibilité de l'environnement autour de l'éolienne. De fait, elles réduisent sans doute les capacités de détection/classification des dispositifs optiques mais aussi radar (la pluie absorbe les ondes radars réduisant alors les distances de détection). Or, ces conditions affectant négativement la probabilité de détection/classification combinée avant une certaine distance peuvent survenir sur une part non négligeable du temps de fonctionnement de certains parcs éoliens (en fonction de leur implantation géographique). Il est donc crucial de réaliser des tests d'évaluation des performances dans toutes les conditions météorologiques et pas seulement dans des conditions idéales (soleil au zénith, beau temps, etc.), afin de mesurer leurs effets sur la probabilité de détection/classification combinée des systèmes.

L'ensemble des conditions météorologiques sera représenté dans le cadre de ce protocole par trois variables continues :

- La visibilité (qui tient compte de la pluie et du brouillard notamment) : mesure de 0 à 10 000 m (selon le modèle de visibilité utilisé).
- La pluviométrie : mesure de 0 et ∞ mm la quantité de pluie tombée par tranche de 10 min. Cette donnée météorologique peut aussi s'exprimer en L/m².
- Le rayonnement global (qui permet d'estimer la couverture nuageuse) : mesure de l'ensemble des rayonnements solaires (direct et diffus) provenant de toutes les directions sur une surface horizontale, dans l'heure précédent le relevé d'un point géographique donné. Cette donnée s'exprime en J/cm².

- **Eclairage lumineux lors de la détection/classification**

Les couchers et levers de soleil peuvent créer de forts effets de contre-jour pour les SDA optiques. De même, une cible dont la trajectoire est parfaitement alignée avec le soleil est beaucoup plus difficile à détecter/classifier par les dispositifs.

Ces paramètres environnementaux sont à considérer dans l'analyse des performances au moyen de deux variables continues, à savoir :

- La luminosité : mesure de 0 à 120 000 lx à l'aide d'un luxmètre.
- L'angle d'incidence du soleil : mesure de 0 à 90°, calculé grâce à la formule suivante : $i = 90 - h$; avec h la hauteur du soleil au moment de la mesure et qui peut être obtenue avec le package « solrad » et la fonction « Incidence » sur R Studio en renseignant la longitude, la latitude et le décalage d'heure (heure d'été/heure d'hiver).
- Azimut du soleil : mesure de 0 à 360°. Ce paramètre peut être calculé grâce au package « solrad » et la fonction « Azimuth » sur R Studio en renseignant la longitude, la latitude et le décalage d'heure (heure d'été/heure d'hiver).

- **L'angle de vue par le système de détection-réaction**

Dans la majorité des cas, les SDA sont fixés au mât ou à proximité d'une éolienne. Les systèmes optiques et radar qui les composent se trouvent à une hauteur (définie comme la distance entre le sol et le SDA) comprise généralement entre 2 et 10 m. La couverture verticale de chaque appareil va donc permettre de surveiller vers le sommet et vers le pied de l'éolienne, selon un certain angle de vue. L'angle d'ouverture vertical correspondant est généralement compris entre 55° et 77° selon les dispositifs, d'après les informations transmises par les fournisseurs de SDA. Or, selon leur hauteur de vol, les oiseaux ne présenteront pas la même surface alaire au SDA. En effet, il existe une corrélation positive entre l'angle de vue entre le SDA et l'oiseau et la surface visible de l'oiseau. Concrètement, lorsqu'un oiseau vole très haut ou très bas (proche des limites du champ de vision vertical des appareils), la surface alaire visible par le SDA sera plus élevée que lorsque l'oiseau vole proche de l'horizon du système (angle de vue vertical se rapprochant de 0°). L'angle de vue par le SDA influe donc probablement sur les capacités de détection/classification des dispositifs. Il doit donc être déterminé à partir de la hauteur de vol et la distance qui sépare le dispositif de l'oiseau et pris en compte dans les analyses statistiques.

Par ailleurs, certaines espèces cibles effectuent des vols verticaux descendants depuis le sommet de l'éolienne. D'autres réalisent des vols verticaux ascendants, c'est-à-dire depuis le pied de l'éolienne. Par conséquent, les SDA doivent détecter au-dessus et en-dessous des turbines pour réduire efficacement les collisions pour ces espèces (voir WP3-R3 du projet MAPE).

L'angle de vue vertical par le SDA sera pris en compte dans les analyses au moyen d'une variable catégorielle se composant de classes de 30° avec des valeurs comprises entre -75° et 105° (Annexe 1). Il est important de garder à l'esprit que la variable considérée dans les analyses est l'angle de l'oiseau par rapport au SDA et non par rapport à l'observateur. Cette donnée sera obtenue *a posteriori* des jours de terrain à partir des valeurs relevées par le télémètre laser (de même que pour les données d'oiseaux équipés de balise GPS). Les scripts d'analyse de données disponibles sur le site MAPE fournissent les formules nécessaires à ces calculs.

- **L'azimut de détection**

Les oiseaux sont susceptibles d'arriver depuis n'importe quelle direction vers les turbines. Ainsi, le SDA doit avoir des performances de détection/classification équivalentes sur 360° à l'horizontale. Par ailleurs, selon les informations qui nous ont été transmises par les entreprises fournisseurs de systèmes optiques de détection-réaction, l'angle d'ouverture horizontal des caméras est compris entre 45° et 90°. Afin de maximiser la couverture horizontale globale du SDA (c'est-à-dire 360°), il est alors nécessaire que les angles de vue horizontaux de deux caméras successives sur le mât de l'éolienne se chevauchent afin d'assurer la continuité de surveillance sur 360°. Ceci soulève deux questions associées aux capacités de détection/classification des SDA de type optique. La première concerne la manière dont le SDA gère un oiseau qui serait détecté au niveau de cette zone de chevauchement (et donc détecté par deux caméras simultanément). La deuxième s'intéresse à l'importance d'évaluer si ces zones de chevauchement sont suffisantes pour couvrir les zones aveugles propres à chaque caméra, liées à l'effet de bord (perte de visibilité sur les bords du champ d'ouverture de chaque caméra).

L'évaluation des performances selon les azimuts est donc nécessaire. Cette variable continue se mesure de 0 à 360°, grâce à une balise géolocalisatrice ou un télémètre laser. Il est important de garder à l'esprit que la variable considérée dans les analyses est l'azimut de l'oiseau par rapport au SDA et non par rapport à l'observateur. Cette donnée sera obtenue *a posteriori* des jours de terrain à partir des valeurs relevées par le télémètre laser ou la balise GPS. Les scripts d'analyse de données disponibles sur le site MAPE fournissent les formules nécessaires à ces calculs.

2.Acquisition des données sur le terrain

2.1 Choix d'un lieu et d'une période

L'objectif d'obtenir une évaluation standardisée des performances de détection/classification combinée avant une certaine distance des SDA qui soit la plus homogène et neutre possible entre les différents systèmes devrait passer par la conduite des tests sur un site commun. Ceci implique un certain nombre de

contraintes à la fois pour les personnes en charge de l'animation de l'évaluation et pour les fournisseurs de SDA, à savoir : un accès au foncier pour installer et sécuriser ces systèmes coûteux ainsi qu'un accès à l'électricité et à internet. Il faut, par ailleurs, que ce site soit occupé ou au moins traversé par un certain nombre d'oiseaux des différentes espèces potentiellement cibles et des trois classes de taille à tester (box 2). A noter qu'il n'est pas nécessaire que ce site soit équipé d'éoliennes. Dans ce cas, les systèmes qui se fixent sur les mâts des éoliennes devraient alors être installés sur des mâts spécifiquement déployés pour ces expérimentations. La mise en place d'un tel site expérimental est compatible avec l'équipement d'oiseaux avec des balises GPS dans le cadre d'un programme de recherche spécifique et permettrait de collecter de manière passive un grand nombre de données nécessaires à l'évaluation des performances des SDA. A la date de rédaction de ce protocole (mai 2023), un tel site n'est pas défini et ce sujet reste ouvert.

En l'absence de site expérimental commun pour mener l'ensemble des tests, chaque fournisseur pourra évaluer son SDA sur un parc déjà équipé. Cette solution ne permet cependant pas d'assurer une équité de traitement entre les différents systèmes, notamment car les conditions environnementales (topographie, arrière-plan, météorologie, espèces d'oiseaux présentes) pourraient ne pas être identiques.

Dans tous les cas, le(s) site(s) choisi(s) nécessite(nt) :

- La présence d'oiseaux sauvages de tailles différentes (box 2 pour les gammes de taille d'espèces cibles)
- La présence éventuelle d'oiseaux sauvages équipés de balises enregistrant une localisation au moins toutes les 5 s dans la zone test. Ceci permettrait d'augmenter de façon importante la taille de l'échantillon sans effort ni coût supplémentaire (voir partie 1.2.)
- Un accès facile (carrossable) avec électricité sur place et foncier sécurisé

La question de la temporalité concernant la mise en œuvre du protocole « fournisseur » ne se pose que lorsque le moyen-tiers mobilisé est la paire d'observateurs humains équipée d'un télémètre laser. En effet, les oiseaux sauvages équipés de balises GPS permettent d'acquérir de façon exhaustive et passive un grand nombre de données tout au long de l'année (ou tout du moins tout au long de leur période d'utilisation du site). De fait, toutes les trajectoires enregistrées à proximité du SDA évalué pourront être analysées, qu'importe la période de l'année.

Pour les relevés de terrain nécessitant la présence d'observateurs humains sur site, il est important de choisir des périodes/moments de l'année qui vont permettre de rencontrer :

- Un panel diversifié de conditions de visibilité et météorologiques (pluie, brouillard, etc.)
- Des oiseaux avec des comportements et vitesses de vol différents. Des mesures seront alors réalisées durant 1) la migration (traversées plutôt rapides et rectilignes de la zone d'étude) et 2) la

période de reproduction (déplacements nombreux, composés de plusieurs changements de direction et des vols pouvant être très lents voire stationnaires pour certaines espèces en chasse).

2.2 L'installation des systèmes de détection-réaction

Certains critères devront être respectés lors de l'installation du SDA sur le site test :

- L'installation du SDA sur le lieu de référence choisi devra être identique à celle réalisée dans un parc éolien. La hauteur à laquelle sont installés les différents systèmes optiques et radar sera mesurée et notifiée par le fournisseur aux personnes qui seront en charge de traiter les données acquises sur le terrain (pour permettre de calculer les valeurs des variables d'influence acquises par un observateur humain, tel que l'angle de vue par exemple).
- Les paramétrages du système seront réalisés par l'entreprise fournisseur selon la méthode de son choix (utilisation de drones, oiseaux sauvages, etc.) et devront correspondre à ce qui est usuellement mis en œuvre lors de l'installation d'un dispositif dans un parc éolien.
- Le dispositif testé dans le cadre de ce protocole devra être identique (hardware, software et composants technologiques) à la version commercialisée au même moment pour les parcs éoliens. La mise sur le marché d'un nouveau modèle ou toute modification significative de la version actuelle imposera de réaliser une nouvelle évaluation. De fait, chaque modèle/version nécessitera l'établissement d'une grille d'évaluation des performances qui lui sera propre (Annexe 1).
- Aucune calibration ou programmation particulière ne doit être réalisée pour « booster » artificiellement les performances des SDA dans le cadre de la mise en œuvre de ce protocole. Notamment, les dispositifs seront configurés pour détecter/classifier strictement les mêmes espèces cibles que celles proposées par le système commercialisé. Une analyse des images de détection après classification sera réalisée afin de donner une indication sur le taux de faux positifs que génère le SDA (box 3) et mis en regard de la probabilité de détection/classification estimée.

Box 3 – Protocole fournisseur - Le nombre de faux positifs pour un système optique de détection-réaction

Le taux (ou nombre) de faux positifs correspond aux événements de détection/classification de cibles qui ne sont pas des oiseaux, ou qui sont des oiseaux mais dans des situations qui ne sont pas à risque, et qui sont pourtant considérées comme tel, engendrant une réaction inutile du système dans le contexte de la réduction des mortalités aviaires. Il existe une corrélation négative entre le nombre de faux positifs et de faux négatifs. En effet, un système qui est paramétré pour maximiser la détection/classification d'oiseaux d'espèces cibles (faible taux de faux négatifs) conduit irrémédiablement à une augmentation de la détection d'objets dont certains sont non-pertinents dans le cadre de la problématique de réduction des mortalités aviaires (fort taux de faux positifs). Il peut donc être intéressant de fournir aux exploitants de parcs le taux de faux positif (entraînant une perte de production inutile) correspondant aux réglages par défaut du SDA.

Ces événements de faux positifs peuvent être évalués *a posteriori* par visionnage des clichés/vidéos enregistrés par le SDA lors des événements détectés/classifiés à risque de collision. Il convient de mobiliser un observateur humain expérimenté en ornithologie afin d'identifier – à l'espèce près – la cible volante relevée par le système : un 0 sera attribué si l'objet volant correspond à un individu d'une espèce d'intérêt ; un 1 sera noté si l'objet volant ne correspond effectivement pas à une espèce ciblée par arrêté préfectoral sur ce parc éolien (= erreur de détection/classification du système). Il s'agit donc d'une démarche particulièrement chronophage.

A noter qu'à ce stade, aucune méthodologie n'a encore été définie pour évaluer ce nombre de faux positifs avec des SDA mobilisant la technologie radar.

- Le système doit être réglé à l'heure locale du réseau (téléphonique ou internet) de façon très précise. L'horodatage du SDA et de l'ensemble des appareils de mesure sera vérifié avant la réalisation des tests sur le terrain pour s'assurer de la bonne analyse des données *a posteriori* (voir la partie 3. ci-dessous).
- La position géolocalisée (3D) précise du SDA (notamment les différentes caméras qui composent les dispositifs optiques) sera relevée en se plaçant à proximité du dispositif ou au pied de l'éolienne selon son lieu de fixation (à proximité ou sur le mât de l'éolienne respectivement).

Les SDA fonctionnent de façon parfaitement autonome. Les données acquises par ces systèmes seront mises à disposition des personnes en charge de l'analyse des données par l'opérateur ou le fournisseur selon les modalités convenues. A noter que selon les dispositifs et les technologies mobilisées, les

informations sauvegardées ne sont pas de même nature (certains indiquent l'espèce ; d'autres, la taille de l'oiseau détecté/classifié ; certains fournissent une mesure de distance à l'oiseau ; d'autres ne fournissent pas cette mesure de distance, etc.).

Par ailleurs, les données extraites des différents dispositifs ne se présentent pas sous le même format (certains fournissent des images ; d'autres, des vidéos ; d'autres, des tableaux au format « csv »).

2.3 Mesure des variables d'influence environnementales

2.3.1. Définition des conditions météorologiques et de l'éclairement lumineux

D'après ce que nous avons décrit précédemment (partie 1.3.), il est essentiel de relever au moins une fois par heure, certaines variables environnementales lors de l'évaluation des performances des SDA, afin de tenir compte de ces variables lors de l'analyse statistique des données. Plus précisément, la visibilité, la luminosité, le rayonnement solaire et la pluviométrie devront être mesurés régulièrement pour que leur influence sur la probabilité de détection/classification combinée avant une certaine distance des SDA soit testée (et rapportée dans la grille de performances).

Les données seront obtenues soit :

- 1) via des instruments de mesure calibrés et installés au plus près du système à tester (box 4). Ces instruments pourront être amenés sur le site d'étude par les observateurs humains équipés d'un télémètre laser pour les jours de test. Une station météorologique permanente, relevant en continu ces données, devra être installée pour toute la durée de l'étude lorsque des oiseaux sauvages équipés de GPS sont impliqués.
- 2) via la PubliTech de MétéoFrance – en y récupérant les données enregistrées par la station météorologique la plus proche du parc éolien sur lequel des tests de performance sont en cours. Cette méthode est pertinente si la station météorologique est à moins de 5 km du site d'étude s'il y a du relief ou à moins de 20 km si le site d'étude se trouve en plaine. Un site se trouvant sur un terrain avec du relief peut subir des conditions météorologiques différentes à une échelle spatiale de quelques kilomètres. Pour assurer la correspondance des données enregistrées par la station et les conditions réellement expérimentées sur site, il convient d'enregistrer les variables environnementales au plus proche du site et donc de ne pas s'éloigner de plus de 5 km. En plaine, les conditions météorologiques sont relativement homogènes dans l'espace, ainsi il est possible d'utiliser des variables environnementales qui auraient été enregistrée jusqu'à 20 km d'un site. Si ce critère de proximité entre site et station météorologique n'est pas rempli, il faudra alors obligatoirement collecter ces données environnementales avec la première méthode, c'est-à-dire grâce à l'utilisation d'instruments de mesure calibrés et installés au plus près du système à tester.

Ainsi, il convient d'anticiper la disponibilité des données fournies par MétéoFrance avant de commencer le terrain afin de prévoir s'il est nécessaire d'échantillonner avec des instruments de mesures et donc d'avoir le temps d'acquérir ce matériel (box 4).

Ces données doivent être obtenues/enregistrées à une fréquence minimale d'une fois par heure.

Box 4 – Instruments de mesure requis pour la définition des conditions environnementales

- Installation d'un visibilimètre d'une portée de 2 km et d'une précision au m (exemple, modèle PWD10 de Vaisala ou CS120A de Campbell Scientific)
- Installation d'un luxmètre d'une précision au lx (exemple, modèle IHM Moineau Instrument 3809SI)
- Installation d'un solarimètre d'une précision au W/m² (exemple, modèle IHM Moineau Instrument 1307SI)
- Installation d'un pluviomètre d'une précision au mm ou L/m² (exemple, modèle Bresser 7002530)

2.3.2. Définition de l'arrière-plan

L'arrière-plan est une variable catégorielle qui prend deux valeurs : « fond de ciel » ou « fond de végétation ». La variable est caractérisée *a posteriori* grâce à un script R (téléchargeable depuis le site de MAPE) qui utilise l'alignement entre la position de l'oiseau et celle du SDA.

A noter que le télémètre a des difficultés pour prendre une mesure sur un fond de végétation. En effet, la coordonnée géographique 3D enregistrée est souvent celle de l'arrière-plan lorsque l'oiseau se trouve sur un fond de végétation. Lorsque cela se produit, il faut alors supprimer ces données (voir paragraphe 2.4.2.).

2.4 Acquisition des données des cibles via un moyen-tiers

2.4.1. Utilisation d'individus sauvages équipés de balises GPS

Si l'étude des performances repose sur l'utilisation d'oiseaux équipés de GPS, les données enregistrées seront accessibles uniquement sur volonté des porteurs de programmes de baguage correspondants. Il conviendra de récupérer les informations suivantes : la date (format : « jj/mm/aaaa »), l'heure (format :

« hh:mm:ss », l'horodatage réseau doit être utilisé pour être strictement identique à celui des autres appareils de mesure et ainsi permettre une mise en correspondance plus aisée des différents jeux de données), la latitude (format WGS84 : degrés décimaux), la longitude (format WGS84 : degrés décimaux), l'altitude (format : m) et l'espèce (et la classe de taille correspondante). A noter que certains modèles de balises GPS permettent de définir des zones (appelées « géofences ») dans lesquelles le relevé de positions 3D se fait en haute résolution (c'est-à-dire un point enregistré toutes les secondes ou toutes les trois secondes) à l'intérieur de ces dernières ; et en basse résolution (c'est-à-dire un point enregistré toutes les XX minutes ou heures) à l'extérieur. Dans le cadre de notre protocole, il conviendra de définir cette géofence avec un rayon de 1 000 m autour du SDA pour acquérir les données les plus précises possibles sur notre zone d'intérêt.

2.4.2. Utilisation d'observateurs humains équipée d'un télémètre laser

Il convient de souligner un certain nombre de prérequis pour que cette approche expérimentale puisse être considérée comme robuste :

- Les deux observateurs humains devront être expérimentés, avertis et formés pour la reconnaissance des espèces cibles rencontrées sur le terrain. Un observateur humain est considéré comme expérimenté quand il a l'habitude des comptages et/ou des inventaires d'oiseaux, notamment en conditions difficiles et durant plusieurs heures consécutives. Il doit également être capable d'identifier en routine et à de grandes distances l'ensemble des espèces cibles fréquentant possiblement le site (et *a fortiori* l'espèce cible du SDA).
- Les deux observateurs devront également avoir participé à une réunion préparatoire pour recevoir les instructions sur le déroulement du protocole (idéalement par l'animateur de la démarche, si un tel animateur existe dans le futur).
- Des jumelles à télémètre laser intégré (instrument de mesure parfois appelé aussi ornithodolite) devront être utilisées pour relever les géolocalisations 3D (longitude, latitude, altitude) des oiseaux tout au long de leurs trajets autour des éoliennes. Le modèle communément utilisé dans les études scientifiques est le Safran® Vectronix Vector 21 (Borkenhagen et al., 2018, Darras et al., 2018, Cole et al., 2019, Therkildsen et al., 2021).
- Les deux observateurs humains et le télémètre laser seront positionnés à une distance suffisante du SDA à surveiller (entre 50 m et 300 m) et à au moins 10 m des voitures. Cela permettra, d'une part, d'optimiser le champ de vision des observateurs et, d'autre part, d'éviter les interférences électromagnétiques par les éoliennes, le SDA ou les voitures sur le télémètre, pouvant générer d'éventuelles erreurs dans les mesures réalisées. A noter que les câbles souterrains peuvent aussi créer des interférences électromagnétiques et qu'il n'est pas possible de connaître leur localisation à l'avance. Il sera donc parfois nécessaire de déplacer les points d'observation si la précision des

mesures du télémètre s'avère trop faible. Une fois le point d'observation choisi, celui-ci pourra être fixé pour l'ensemble des jours de terrain.

- Le télémètre laser sera calibré avant son utilisation, au moins au début de chaque semaine de relevés, pour assurer la précision de mesure.
- Le degré de déclinaison magnétique sera corrigé sur le télémètre laser au début de chaque semaine de relevés.
- Le relevé de trajets se fera sur une sphère de 1 000 m de rayon par rapport au SDA surveillé.
- La sauvegarde des données relevées par l'ornithodolite se fera sur une tablette, directement sur le terrain grâce à l'application LRF DataViewer (même en présence de pluie). Les données enregistrées sur l'application au fur et à mesure de la journée devront être exportées toutes les heures au format « .csv » afin de créer des sauvegardes sur la tablette. Chaque position sera automatiquement associée à un identifiant chiffré (attribué par ordre croissant) et sera reportée sur une carte avec fond orthophotographique en temps réel au moyen d'une icône et du chiffre correspondant. Afin de retrouver les points correspondants au même trajet, l'observateur humain ayant le rôle de scribe (voir ci-dessous pour la répartition des rôles) devra renommer le premier et dernier point du trajet avec le même identifiant.
- Un horodatage de la tablette, identique à celui du SDA, devra impérativement être effectué avant le relevé des premières mesures. La similitude des fuseaux horaires entre les jumelles à télémètre laser intégré et le SDA sera vérifiée en se connectant à l'interface de contrôle du dispositif de réduction des mortalités aviaires. Le cas échéant, il sera déterminé le décalage horaire qu'il existe entre les deux horodatages afin de pouvoir faire correspondre les données enregistrées par ces deux appareils (SDA et jumelles à télémètre laser intégré) *a posteriori*.

Le « Guide d'accompagnement au déploiement des protocoles d'évaluation des performances des systèmes de détection-réaction » disponible sur le site MAPE, permet d'obtenir plus de précision sur les points mentionnés ci-dessus et notamment concernant le choix du point d'observation, le calibrage du télémètre laser, la correction du degré de déclinaison magnétique, l'utilisation de l'application LRF DataViewer ou encore le contrôle de l'horodatage.

Les deux observateurs humains scruteront le ciel à l'aide de jumelles binoculaires ou longues-vues, pour détecter puis identifier les espèces. Ils réaliseront ces observations sur une plage de six heures consécutives selon un schéma standardisé. Ils enchaîneront six fois de suite la séquence suivante : 45 min d'observation ; 15 min de pause. Cette durée doit permettre d'optimiser le temps d'observation par journée de terrain tout en assurant une attention maximale des deux observateurs sur toute la durée de la période de récolte de données. Afin de prendre en compte une diversité de paramètres abiotiques et biotiques représentatifs des conditions de fréquentation du site par les oiseaux, l'heure de début de la journée de six heures d'observation sera différente d'un jour à l'autre (tableau 1). Parmi les différents horaires de début et

de fin d'observation, il faudra des journées qui commencent au lever du soleil et d'autres qui finissent au coucher du soleil.

Tableau 1 : Exemple de plages horaires de six heures consécutives pour réaliser la récolte des données lors des jours de terrain par la paire d'observateurs humains équipée d'un télémètre laser

Jour 1	Jour 2	Jour 3	Jour 4	Jour 5	Jour 6	Jour 7	Jour 8	Jour 9	Jour 10
7H30 – 13H30	12H30 – 18H30	9H00 – 15H00	14H00 – 20H00	10H30 – 16H30	13H00 – 19H00	8H00 – 14H00	14H30 – 20H30	8H30 – 14H30	10H00 – 16H00

Lorsqu'un oiseau s'approchera de la zone d'intérêt (i.e. sphère de 1 000 m de rayon autour de l'éolienne), l'un des observateurs suivra l'oiseau à l'aide du télémètre laser et deviendra alors collecteur de données. Il relèvera toutes les 5 s (délai d'acquisition entre deux points successifs pour un observateur humain entraîné à l'utilisation du télémètre laser) la position de l'oiseau, tout au long de sa trajectoire.

Le second observateur humain sera le scruteur de ciel. Il assurera à l'œil nu (pour maximiser le champ de vision surveillé) une veille des environs du dispositif pour avertir le collecteur de données qu'un autre oiseau intéressant à suivre est disponible (i.e. individu à 1 000 m ou moins de l'éolienne) une fois que l'oiseau en cours de suivi par le télémètre n'est plus considéré comme pertinent (car s'éloignant de la zone à risque par exemple). **A noter qu'il est très important pour la robustesse des analyses de privilégier le suivi d'une longue trajectoire pour quelques individus plutôt que de courts fragments de trajets pour beaucoup d'oiseaux.**

Le scruteur de ciel s'occupera également de prendre en note toutes les informations énoncées par le collecteur de données : début et fin d'un trajet (relevé des identifiants des points associés) ; confirmation de l'espèce de l'oiseau ciblé avec le télémètre laser ou *a minima* sa classe de taille, suppression immédiate des potentielles erreurs de localisations/distances relevées au télémètre (aisément identifiable selon les distances mesurées juste avant et juste après ou en regardant la position du point sur la cartographie de l'application LRF DataViewer).

Le scruteur de ciel devra relever l'information de la présence simultanée de plusieurs oiseaux sur le terrain (un seul oiseau ou plusieurs oiseaux). Certains SDA n'ont pas la possibilité d'indiquer la position de l'oiseau détecté (azimut ou position GPS). Ainsi, pour ces systèmes, lors des analyses, il faudra supprimer les trajectoires pour lesquelles plusieurs oiseaux étaient présents simultanément sur le terrain car il ne sera pas possible de déterminer lequel était concerné par la détection du SDA.

Les observateurs peuvent échanger leur rôle au cours d'une semaine de relevés, tant que l'observateur au télémètre est formé à son utilisation. Au cours d'une journée, les rôles devront, dans la mesure du possible, être conservés. Un modèle de fiche terrain est disponible dans le « Guide d'accompagnement au déploiement des protocoles d'évaluation des performances des systèmes de détection-réaction » disponible sur le site MAPE.

Il est important de noter que pour ce protocole « fournisseur », les observateurs devront s'efforcer, si les conditions de terrain le permettent, de balayer les gradients de variables décrites ci-dessus (gamme de taille des espèces, différents arrière-plans, etc.) de manière équilibrée pour assurer la robustesse des analyses statistiques. Une veille sur les données déjà acquises lors des sessions de terrain précédentes sera donc nécessaire pour essayer de compléter au mieux les relevés sur des situations moins fréquemment représentées dans le jeu de données.

Les tests réalisés sur le terrain en 2022 ont montré qu'il existait un certain nombre de situations où les observateurs avec télémètre et le SDA n'avaient pas des probabilités de détection indépendantes (Ballester et al., 2023). Cette hypothèse d'indépendance est pourtant importante/primordiale pour que les estimations des performances soient non-biaisées. Malheureusement, aucune solution n'existe pour remédier aux quelques cas identifiés qui conduisent tous à une légère sur-estimation des performances en particulier pour le protocole « parc » (box 5). Dans le cadre du protocole « fournisseur », ces situations auront pour conséquence de rendre plus complexe l'estimation précise des performances dans les situations limites (taille d'échantillons plus faibles).

Box 5 – Cas de non-indépendance entre observateurs avec télémètre et SDA

L'utilisation du télémètre laser a posé problème dans certaines conditions lors de la phase de test des protocoles sur le terrain courant 2022, problèmes qui ont des conséquences sur l'estimation des performances de détection/classification des SDA.

Ainsi, lors du suivi de trajectoires d'oiseaux sur fond de végétation, la distance relevée par le télémètre est souvent celle entre le télémètre et l'arrière-plan et non celle entre le télémètre et l'oiseau (qui est pourtant celle qui nous intéresse). Ces erreurs induisent un biais puisqu'il est alors difficile d'enregistrer les trajectoires des oiseaux volant au ras du sol ou à proximité des zones de végétation denses. Or, ces configurations sont aussi probablement plus difficiles à détecter par les SDA.

En présence de pluie et de vent ou d'une forte pluie, il n'est pas possible de protéger correctement le télémètre laser tout en poursuivant les relevés de trajectoires. Etant un appareil de mesure coûteux, la récolte de données ne peut donc être réalisée dans de telles conditions météorologiques. La prise de mesure en condition de brouillard n'est par ailleurs pas possible. En effet, le laser est réfléchi par les gouttes qui composent le brouillard, générant des distances ne correspondant pas à celles entre le télémètre et les objets/cibles. Ainsi, les possibilités d'effectuer des relevés en condition de pluie et de brouillard sont limitées. Les conditions très défavorables à la préservation du matériel ou à la prise de mesure sont probablement aussi celles limitantes pour la détection des SDA.

Ainsi, les capacités de détection avec le télémètre laser ne sont pas totalement indépendantes de celles des systèmes. L'approche de modélisation retenue pose pourtant l'hypothèse d'une indépendance de la détection entre observateur et système qui est ici violée pouvant de fait conduire à une estimation biaisée des performances. Dans les deux situations ci-dessus, le biais conduit à surestimer les capacités de détection des systèmes puisque les observateurs ont moins de trajectoires dans des conditions de potentielle faible détection des SDA dans les données que ce qu'il y a dans la réalité. Ce problème est particulièrement vrai pour le protocole « parc » mais limité pour le protocole « fournisseur » qui balaie les différentes conditions de détection. Cependant cela peut conduire à avoir des tailles d'échantillons faibles dans certaines modalités des variables d'influence et donc des estimations imprécises dans ces conditions difficiles.

Enfin, en présence de plusieurs oiseaux, pour les SDA n'ayant pas la possibilité de collecter une information sur la position géographique ou l'azimut de l'oiseau détecté, il est alors impossible de savoir si l'oiseau suivi avec le télémètre correspond à celui détecté par le système. La solution pour ces systèmes est alors de noter les trajectoires concernées par la présence de groupe d'oiseaux dans la zone à risque afin de pouvoir supprimer ces trajectoires lors des analyses de données. Ceci permet de limiter les biais.

3. Traitement et modélisation statistique des données acquises sur le terrain

L'estimation de la probabilité de détection/classification combinée du système avant une certaine distance correspond à la proportion de situations où le système a détecté/classifié un oiseau d'une espèce d'intérêt avec une trajectoire à risque parmi l'ensemble des trajets d'oiseaux équipés de GPS ou ceux ayant été détectés/classifiés comme tels par la paire d'observateurs humains équipée d'un télémètre laser et cela avant une certaine distance.

Pour obtenir ces données, il convient de mettre en correspondance les trajectoires relevées par ces observateurs (grâce aux jumelles avec télémètre laser intégré) ou les trajectoires des oiseaux équipés de balise GPS dans les 1 000 m autour du SDA avec les données collectées par ce dernier dans cette même zone. A noter que dans le cadre de cette procédure, nous excluons du jeu de données à analyser les trajets ou portions de trajets qui s'éloignent du SDA. En effet, ces situations ne présentent pas un risque de collision pour les oiseaux et ne présentent donc aucun intérêt pour cette évaluation de performance. Ces trajets ou portions de trajets sont identifiés avant analyse en utilisant les localisations des points successifs.

Concrètement, la distance entre l'oiseau et le SDA sera calculée via des formules utilisant les positions géographiques de ces derniers (voir les guides des scripts d'analyses de données). En effet, au vu des prérequis détaillés dans la partie 2.4.2., ces derniers ne seront pas exactement au même endroit que le SDA. Cela nécessitera de recalculer la distance oiseau-SDA en utilisant les coordonnées 3D de l'oiseau (enregistré par le télémètre ou par une balise GPS pour les oiseaux équipés) et du SDA (fournies par l'opérateur). De plus, il est nécessaire de mesurer une distance réelle entre l'oiseau et le SDA et non pas une simple distance projetée sur le sol afin d'intégrer le fait que l'observateur et le SDA ne sont pas nécessairement à la même hauteur.

A noter que l'altitude mesurée par le télémètre et les balises GPS posées sur les oiseaux est l'altitude par rapport au niveau de la mer. Avec ces deux approches, il faudra donc utiliser l'altitude du SDA par rapport au niveau de la mer. Cette dernière n'est généralement pas donnée par les opérateurs/fournisseurs, seulement la hauteur du système peut être obtenue. Il faudra alors corriger l'altitude du SDA en additionnant sa hauteur à celle de l'altitude de la localisation fournie par un MNT (Modèle Numérique de Terrain).

Pour réaliser la mise en correspondance entre les deux jeux de données (moyen-tiers *versus* SDA), il convient de rechercher, parmi les données enregistrées par le SDA, le moment où le trajet relevé au télémètre ou enregistré par la balise GPS de l'oiseau sauvage est détecté pour la première fois par le SDA. Ceci doit se faire par la mise en relation des heures de détection/classification. Il est donc crucial que l'horodatage du télémètre laser et du GPS des oiseaux équipés soit strictement identique à celui du SDA. Il

est important de rappeler ici que le télémètre laser enregistre une localisation toutes les 5 secondes environ, tandis que le SDA en collecte en continu. La correspondance entre les deux jeux de données se fait donc via l'identification du relevé du SDA le plus proche temporellement de celui du télémètre laser ou de la balise GPS. Une tolérance de plus ou moins 20 secondes est acceptée. Cette tolérance peut être modifiée dans les scripts d'analyse de données.

Une fois que cette recherche est accomplie, le plus simple pour constituer le jeu de données à analyser est de partir du fichier brut regroupant les géolocalisations 3D collectées par le télémètre (ou celui issu du GPS des oiseaux équipés, dans le cas d'un suivi passif par balises géolocalisatrices), après exclusion des portions de trajets s'éloignant du SDA (tableau 2). Concrètement, une colonne renseignant sur la détection du trajet ou non par le dispositif de réduction des mortalités aviaires sera complétée/ajoutée : un 0 sera attribué avant le point de correspondance entre relevé au télémètre/trajet GPS et l'évènement de détection/classification par le système et un 1 sera attribué après cette correspondance. Cette procédure implique de considérer que si le système a détecté un trajet à une certaine distance, alors ce trajet est considéré comme détecté à des distances inférieures.

A noter que certains trajets relevés par les observateurs/par le GPS d'un oiseau équipé peuvent ne pas du tout être détectés par le SDA. Dans ce cas, la colonne « détection » contiendra uniquement des 0 pour l'ensemble des points d'un même trajet qui ont une combinaison unique de classes de variables environnementales. A l'inverse, certains trajets peuvent être détectés par le SDA mais pas par les observateurs. Dans ces cas-là, ces trajets ne sont pas utilisés (box 1).

Par la suite, pour chaque point relevé au télémètre laser (identifié dans les analyses grâce à un « ID_unique » qui lui est propre) ou par la balise GPS d'un oiseau équipé, le fichier brut sera enrichi des données collectées pour toutes les variables explicatives ayant une influence potentielle sur la probabilité de détection/classification combinée avant une certaine distance, à savoir : la classe de taille de l'espèce, l'arrière-plan, la visibilité, la pluviométrie, le rayonnement solaire, la luminosité et l'angle d'incidence du soleil (voir partie 1.3.).

Pour les informations environnementales obtenues au moyen d'appareils de mesure, lorsqu'aucune donnée n'a été relevée exactement au moment même d'une géolocalisation (par un observateur humain équipé d'un télémètre laser ou un GPS placé sur un oiseau), la donnée environnementale la plus proche temporellement de la localisation 3D sera conservée.

Après exclusion des éventuelles erreurs de mesure, l'angle de vue par le dispositif et l'azimut entre le SDA et l'oiseau cible seront calculés grâce aux coordonnées GPS 3D de ces derniers.

Le fichier de données ainsi constitué se présente alors comme suit dans le tableau 2 :

Tableau 2 : Jeu de données fictif à analyser, c'est-à-dire après mise en correspondance des fichiers bruts issus des relevés au télémètre laser/des balises GPS et des enregistrements réalisés par le SDA

Id_trajet	TimeStamp	Detection par le SDA	Classe distance oiseau-SDA	Classe angle oiseau	Classe azimuth oiseau	Présence d'un ou plusieurs oiseaux	Description	Incidence du soleil	Azimut du soleil	Pluie	Rayonnement global	Visibilité	Arrière plan	Luminosité
12	22/08/2022 12:19	0	4	3	5	2	Vfauve	57,22	197,86	0	166	NA	ciel	NA
12	22/08/2022 12:20	0	3	3	5	2	Vfauve	57,21	198,06	0	166	NA	ciel	NA
12	22/08/2022 12:22	0	3	2	5	2	Vfauve	57,13	199,29	0	166	NA	ciel	NA
12	22/08/2022 12:23	1	3	1	5	2	Vfauve	57,12	199,54	0	166	NA	ciel	NA
12	22/08/2022 12:20	0	2	2	5	2	Vfauve	57,19	198,41	0	166	NA	ciel	NA
12	22/08/2022 12:20	0	2	3	5	2	Vfauve	57,19	198,33	0	166	NA	ciel	NA
13	22/08/2022 12:25	0	4	2	6	2	Vfauve	57,05	200,44	0	166	NA	vegetation	NA
13	22/08/2022 12:25	0	3	2	6	2	Vfauve	57,04	200,62	0	166	NA	vegetation	NA
13	22/08/2022 12:23	0	2	3	4	2	Vfauve	57,11	199,61	0	166	NA	vegetation	NA
13	22/08/2022 12:23	0	1	2	5	2	vfauve	57,10	199,72	0	166	NA	ciel	NA
14	22/08/2022 12:31	0	4	3	1	2	cjeanblanc	56,83	203,10	0	326	NA	ciel	NA
14	22/08/2022 12:30	1	3	3	6	2	cjeanblanc	56,88	202,57	0	326	NA	ciel	NA
14	22/08/2022 12:30	1	3	3	1	2	cjeanblanc	56,87	202,63	0	326	NA	ciel	NA
14	22/08/2022 12:29	0	3	3	5	2	cjeanblanc	56,91	202,23	0	166	NA	ciel	NA
14	22/08/2022 12:29	1	3	3	5	2	cjeanblanc	56,91	202,29	0	166	NA	ciel	NA
14	22/08/2022 12:28	0	2	3	5	2	cjeanblanc	56,96	201,64	0	166	NA	ciel	NA
14	22/08/2022 12:27	0	1	3	5	2	cjeanblanc	56,99	201,29	0	166	NA	ciel	NA

L'analyse de ce jeu de données (tableau 2) pour obtenir des probabilités de détection/classification combinée avant une certaine distance ainsi que la précision d'estimation associée se réalise au moyen d'un Modèle Linéaire Généralisé (GLM) avec une distribution binomiale des données et un lien *logit*. Les données collectées sur un même trajet n'étant pas indépendantes, un effet aléatoire sur la variable « identifiant du trajet » devrait être inclus dans le modèle pour gérer la pseudo-réplication des données. Cependant, ce modèle mixte ne converge pas. En effet, la modélisation de la capacité à détecter avant une certaine distance implique que toutes les lignes avec des distances plus proches qui suivent temporellement la détection sont considérées comme détectées (valeur renseignée = « 1 »). Ceci entraîne une autocorrélation temporelle presque parfaite, ne permettant pas la gestion de la pseudo-réplication par un effet aléatoire. La solution pour remédier au problème de pseudo-réplication est de ne plus considérer qu'une seule unité statistique par trajectoire (une seule ligne du tableau 2 par trajectoire). Cependant, au cours d'un même trajet d'oiseau, son azimuth, son angle vertical et sa distance au SDA vont varier. Ce sont ces mêmes variations de position qui peuvent influencer la détection et ne garder qu'une position aléatoire par trajectoire réduit de fait très fortement la puissance statistique. L'approche retenue consiste à conserver une seule localisation par modalité combinée de distance/azimut/angle et ces données sont analysées avec un simple GLM incluant ces variables comme variables explicatives.

Les trajets relevés sur différents jours sont analysés simultanément au moyen d'un jeu de données unique. Il est possible de réaliser des analyses sur des sous-ensembles de données dans le cas où l'on ne souhaite pas remplir certaines parties de la grille de performances. Par exemple, si des trajets de « petites » espèces sont relevés mais que le fournisseur considère que son système n'est pas adapté à ces espèces, dans ce cas, il peut soit réaliser les analyses pour montrer que la probabilité de détection/classification combinée avant une certaine distance pour ces espèces est faible, soit supprimer les données de ces petites espèces et mettre des NA dans la grille d'évaluation des performances.

Pour plus de précision sur les analyses et les modèles utilisés, des scripts d'exemple sont disponibles sur le site MAPE.

Les prédictions du modèle sont ensuite produites pour chaque élément de la grille de performances pour obtenir les probabilités prédites de détection/classification combinée en fonction des variables explicatives ainsi que leurs intervalles de confiance à 95%.

4. Tailles d'échantillon et précision des estimations

L'incertitude sur les paramètres estimés est directement liée à la taille de l'échantillon (nombre de fois où la variable à expliquer a été mesurée). Il existe de fait une relation mathématique entre précision d'une

estimation et taille d'échantillon utilisé pour l'estimer. Plus l'échantillon est grand, plus l'incertitude est faible et plus les valeurs des paramètres estimés sont proches de la réalité.

De fait, il est possible d'obtenir des estimations très différentes de la valeur réelle d'un paramètre en se basant sur des échantillons de tailles diverses. Ce biais d'estimation est exacerbé lorsque l'échantillon est de petite taille. En effet, on sait, par le théorème central limite, que plus la taille de l'échantillon augmente, plus la valeur estimée a de chance de se rapprocher de la valeur réelle. A l'inverse, la valeur estimée du paramètre peut être très éloignée de la valeur réelle lorsque l'échantillon est réduit.

Ce théorème central limite est aujourd'hui largement utilisé pour calculer l'intervalle de confiance (à un seuil donné) d'une estimation sachant sa taille (et les caractéristiques de la population statistique étudiée). La figure 2 présente les valeurs estimées de la probabilité de détection/classification combinée avant une certaine distance et les intervalles de confiance associés à 95% attendus pour différentes valeurs réelles du paramètre et différentes tailles d'échantillon (voir aussi l'annexe 2). Ces graphiques montrent clairement que la largeur de l'intervalle de confiance décroît en suivant une forme exponentielle négative, ce qui corrobore la théorie.

Cette figure illustre aussi qu'au-delà d'une certaine taille d'échantillon (associée à un effort de mesure donné), la prise en considération de données supplémentaires ne permet pas d'améliorer sensiblement la précision. Ces figures peuvent être utilisées pour guider sur le nombre de trajets à relever pour obtenir un degré de précision voulu. L'examen de la figure 2 suggère qu'une taille d'échantillon d'environ 100 trajets permet d'obtenir une bonne précision de l'estimation. Le protocole « fournisseur » implique de modéliser l'effet de plusieurs variables explicatives sur la probabilité de détection/classification combinée avant une certaine distance. De fait, l'effort de mesure devra être plus important que cette taille d'échantillon limite pour obtenir une estimation robuste.

A noter que la précision de l'évaluation des capacités de détection dépend elle-même de la probabilité de détection. En effet, une estimation réalisée avec un même nombre de trajets sera moins précise pour un SDA qui a une probabilité de détection plus faible (Annexe 2).

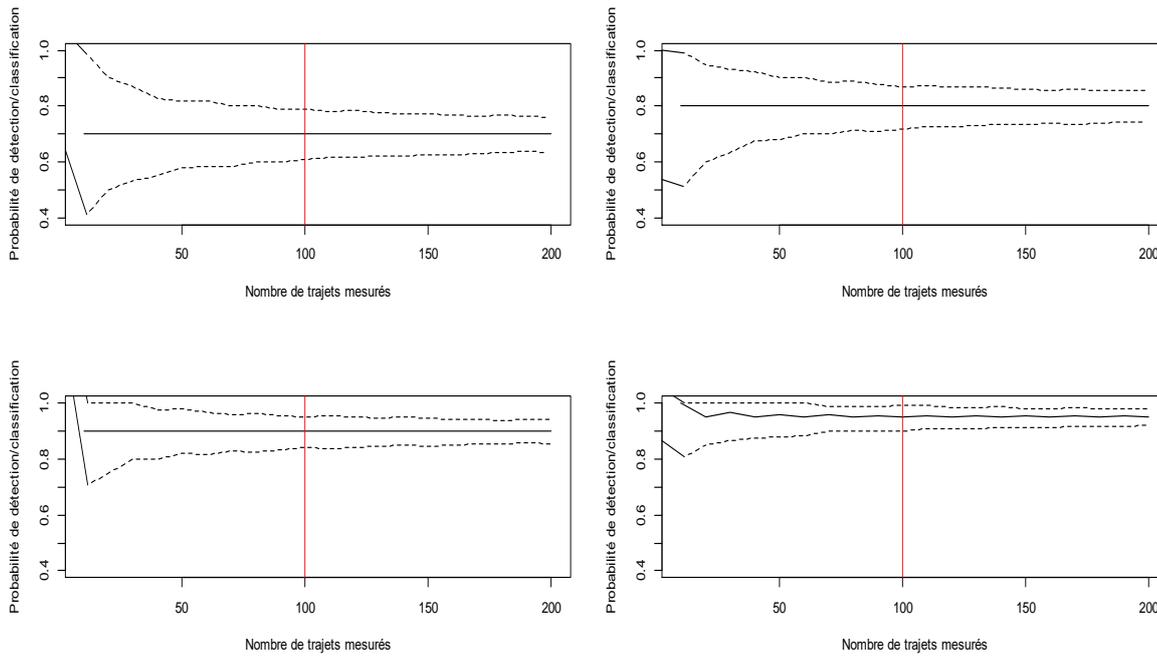


Figure 2 : Estimation moyenne et son intervalle à 95% de la probabilité de détection/classification combinée avant une certaine distance, selon différents efforts de relevé de trajets et différentes valeurs réelles de cette probabilité (0,70 en haut à gauche ; 0,80 en haut à droite ; 0,90 en bas à gauche et 0,95 en bas à droite)

Références

Borkenhagen, K., Corman, A. M., & Garthe, S. (2018). Estimating flight heights of seabirds using optical rangefinders and GPS data loggers: a methodological comparison. *Marine Biology*, 165(1), 1-7.

Cole, E. L., Waggitt, J. J., Hedenstrom, A., Piano, M., Holton, M. D., Börger, L., & Shepard, E. L. (2019). The Ornithodolite as a tool to quantify animal space use and habitat selection: a case study with birds diving in tidal waters. *Integrative zoology*, 14(1), 4-16.

Darras, K., Furnas, B., Fitriawan, I., Mulyani, Y., & Tschardt, T. (2018). Estimating bird detection distances in sound recordings for standardizing detection ranges and distance sampling. *Methods in Ecology and Evolution*, 9(9), 1928-1938.

McClure CJW, Martison L, Allison TD (2018) Automated monitoring for birds in flight_ Proof of concept with eagles at a wind power facility. *Biological Conservation*:9

Therkildsen, O. R., Balsby, T. J., Kjeldsen, J. P., Nielsen, R. D., Bladt, J., & Fox, A. D. (2021). Changes in flight paths of large-bodied birds after construction of large terrestrial wind turbines. *Journal of Environmental Management*, 290, 112647.

Annexe 1 - Modèle de grille finale des performances estimées pour chaque SDA par application du protocole « fournisseur »

Système de détection-réaction évalué :		Oiseaux de petite/moyenne/grande taille		
Distance dispositif - éolienne :		Probabilité de détection	IC inférieur (95%)	IC supérieur (95%)
Distance (m)	0-100			
	100-200			
	200-300			
	300-400			
	400-500			
	500-600			
	600-700			
	700-800			
	800-900			
	900-1000			
Azimut de l'oiseau (°)	0-60			
	60-120			
	120-180			
	180-240			
	240-300			
	300-360			
Angle vertical de l'oiseau (°)	(-75)-(-45)			
	(-45)-(-15)			
	(-15)-15			
	15-45			
	45-75			
Pluviométrie (mm/10min)	0-0,58			
	0,58-1,25			
	> 1,25			
Rayonnement global (J/cm²/1h)	0-70			
	70-140			
	140-210			
	210-280			
	280-350			
Azimut du soleil (°)	0-90			
	90-180			
	180-270			
	270-360			
Incidence du soleil (°)	(-20)-10			
	10-40			
	40-70			

Arrière-plan	Fond de ciel			
	Fond de végétation			
Visibilité (m)	0-200			
	200-400			
	400-600			
	600-800			
	800-1 000			
Luminosité (lx)	0-24 000			
	24 000-48 000			
	48 000-72 000			
	72 000-96 000			
	96 000-120 000			

Annexe 2 - Intervalles de confiance à 95% obtenus en fonction du nombre de trajectoires utilisé et de la probabilité de détection du SDA

Nombre de trajectoires	Probabilité de détection	IC 95% inférieur	IC 95% supérieur
10	0,5	0,2	0,8
20	0,5	0,3	0,7
30	0,5	0,33	0,67
40	0,5	0,35	0,65
50	0,5	0,36	0,64
60	0,5	0,38	0,62
70	0,5	0,39	0,61
80	0,5	0,39	0,61
90	0,5	0,4	0,6
100	0,5	0,4	0,6
110	0,5	0,41	0,59
120	0,5	0,41	0,59
130	0,5	0,42	0,58
140	0,5	0,42	0,59
150	0,5	0,42	0,58
160	0,5	0,42	0,58
170	0,5	0,42	0,58
180	0,5	0,43	0,57
190	0,5	0,43	0,57
200	0,5	0,43	0,57

Nombre de trajectoires	Probabilité de détection	IC 95% inférieur	IC 95% supérieur
10	0,7	0,4	1
20	0,7	0,5	0,9
30	0,7	0,53	0,87
40	0,7	0,55	0,83
50	0,7	0,56	0,82
60	0,7	0,58	0,82
70	0,7	0,59	0,8
80	0,7	0,6	0,8
90	0,7	0,6	0,79
100	0,7	0,61	0,79
110	0,7	0,61	0,78
120	0,7	0,62	0,78
130	0,7	0,62	0,78
140	0,7	0,62	0,78
150	0,7	0,63	0,77
160	0,7	0,63	0,77
170	0,7	0,63	0,77
180	0,7	0,63	0,77
190	0,7	0,63	0,76
200	0,7	0,64	0,77

Nombre de trajectoires	Probabilité de détection	IC 95% inférieur	IC 95% supérieur
10	0,8	0,5	1
20	0,8	0,6	0,95
30	0,8	0,63	0,93
40	0,8	0,68	0,93
50	0,8	0,68	0,9
60	0,8	0,7	0,9
70	0,8	0,7	0,89
80	0,8	0,71	0,89
90	0,8	0,71	0,88
100	0,8	0,72	0,87
110	0,8	0,72	0,87
120	0,8	0,73	0,87
130	0,8	0,73	0,87
140	0,8	0,73	0,86
150	0,8	0,73	0,86
160	0,8	0,74	0,86
170	0,8	0,74	0,86
180	0,8	0,74	0,86
190	0,8	0,74	0,86
200	0,8	0,75	0,86

Nombre de trajectoires	Probabilité de détection	IC 95% inférieur	IC 95% supérieur
10	0,9	0,7	1
20	0,9	0,75	1
30	0,9	0,77	1
40	0,9	0,8	0,98
50	0,9	0,8	0,98
60	0,9	0,82	0,97
70	0,9	0,83	0,96
80	0,9	0,83	0,96
90	0,9	0,83	0,96
100	0,9	0,84	0,96
110	0,9	0,84	0,95
120	0,9	0,84	0,95
130	0,9	0,85	0,95
140	0,9	0,85	0,95
150	0,9	0,85	0,95
160	0,9	0,85	0,94
170	0,9	0,85	0,94
180	0,9	0,86	0,94
190	0,9	0,86	0,94
200	0,9	0,86	0,94

