

# Analyse du territoire et propositions d'aménagements en faveur du Lucane cerf-volant

Programme LIFE Biocorridors -  
14 NAT/FR/000290

Rapport final – décembre 2018

B. Dodelin, G. Fantino - Geopeka, G. Darmon



Biosphärenreservat  
Pfälzerwald-  
Nordvogesen



Dodelin, B., Fantino, G., Darmon, G. 12/2018. *Analyse du territoire et propositions d'aménagements en faveur des espèces saproxyliques et notamment le Lucane cerf-volant. Programme LIFE Biocorridors – 14 NAT/FR/000290*. Rapport d'étude pour le PNR des Vosges du Nord et le Biosphärenreservat Pfälzerwald, 43 pp.

Avec ce rapport, sont rendus plusieurs documents annexes sous différents formats électroniques :

- Les données cartographiques (fichiers QGIS)
- Cartes des variables (images)
- Les corrélations entre variables (tableur)
- Les lignes de code pour le logiciel R (texte et fichier R)
- Les données brutes (fichiers tableurs .csv)

## Table des matières

<b>1. Rappels de contexte et d'objectifs</b> .....	<b>7</b>
<b>2. Méthodes d'analyses</b> .....	<b>8</b>
a) Démarche globale de l'analyse .....	8
b) Données .....	8
(1) Structuration des variables .....	8
(2) Variables environnementales .....	9
(3) Mise à jour des données de Lucane cerf-volant .....	11
(4) Données de coléoptères saproxyliques .....	12
c) Choix des modèles .....	13
(1) RF - Random Forest.....	13
(2) GLM - Generalized Linear Model .....	14
(3) ZI - Zero-Inflated Negative Binomial regression .....	14
(4) Prédiction de distributions.....	14
d) Construction de scénarios .....	15
e) Autres outils statistiques .....	17
(1) Test de Spearman .....	17
(2) Analyse de variance .....	17
f) Préparation des données : corrélations entre variables .....	17
(1) Variables pédologiques .....	17
(2) Variables d'occupation du sol .....	17
(3) Variables de topologie.....	18
(4) Variables climatiques .....	19
<b>3. Résultats</b> .....	<b>21</b>
a) Performances des modèles.....	21
b) Facteurs explicatifs de la distribution du Lucane cerf-volant.....	21
(1) Variables principales selon le modèle RF .....	21
(2) Variables principales selon le modèle GLM .....	22
(3) Variables principales selon le modèle ZI .....	23
c) Prédiction de présence du Lucane cerf-volant .....	24
d) Analyse d'après les prédictions de présence du Lucane .....	27
(1) Prédiction et variables "de gestion".....	27
(2) Prédiction et scénarios.....	28
(3) Prédiction et réseaux d'îlots de sénescence.....	31
<b>4. Conclusions</b> .....	<b>33</b>
<b>5. Bibliographie</b> .....	<b>36</b>
<b>6. Annexe 1. Liste des contacts et réunions de travail</b> .....	<b>38</b>
<b>7. Annexe 2. Carte détaillée des prédictions de Lucane</b> .....	<b>41</b>

## Abstract

This study covers the territory of the Transboundary Biosphere Reserve of the Northern Vosges - Pfälzerwald. The general objective is to determine if this space guarantees the necessary conditions for the existence of saproxylic species and in particular the stag beetle, an emblematic species.

A large dataset has been assembled, gathering 50 environmental and anthropic variables, including forest management, and 424 stag beetle observations. Stag beetle data were covering 242 km<sup>2</sup> among a total of 3 517 km<sup>2</sup> (6,89%). Data comes from online databases, private collections and citizen science programs from both French and German territories. They cover the 1990-2018 period, except for forest continuity which is reflecting 200 years of forest presence.

The analysis was based on three model types: Random Forest (RF), Generalized Linear Model (GLM), Zero-Inflated Negative Binomial regression (ZI), the last one taking into account the presence of false-zero. The factor that best explains the stag beetle distribution was ANTH, the presence of buildings. This variable is 100 time more influential than any other variable. This suggests a serious bias in the observation data. However, this kind of bias is regularly present in all studies using citizen science. As a consequence, ANTH must be seen as a substitute for observers density. Thus, the observation map reflected contacts between species and observers and not the real presence of the stag beetle.

Models and scenarios proposed have shown that it is possible to evaluate which environments are favourable for the stag beetle. Moreover, important applications are possible, although our tests must be considered as examples of evaluation procedures. These applications allow detailed investigations for various management scenarios built to favour the presence of the stag beetle. Furthermore, the models can be used with any distribution dataset, not only the stag beetle. They are strong tools to link species observations with environmental factors, for finally predicting species distributions.

Further studies on stag beetle should focus on a more complete sampling. It was shown that surveys have not sufficiently covered certain habitats and have not been spatially homogeneous. Consequently, too many false negatives (absence of observations) are present in the observation data. Ideally, new observations shall be made at random, on all types of land use and status (protected areas, old-growth forest, etc.). The overall observation rate should be increased to cover about 457 new km<sup>2</sup> if one wanted to reach the observations pressure obtained in Belgium (20,5 %, Thomaes *et al.* 2008). This goal might be hard to achieve, except by focussing on small areas to be intensively investigated. Models can therefore be calibrated on those small areas and then applied to the complete dataset and landscape. New investigations may benefit from the observers network put in place prior to the study and promoted during its development. Small areas should be chosen far from towns (> 2 km) and include protected areas and varied landscape contexts. This would benefit from event-type animations and communication about existing stag beetle observation programs ongoing both French and German sides of the transboundary biosphere reserve.

## Résumé

Cette étude couvre le territoire de la réserve de biosphère transfrontalière des Vosges du Nord - Pfälzerwald. L'objectif général est de déterminer si cet espace garantit les conditions nécessaires à l'existence des espèces saproxyliques et en particulier du lucane cerf-volant, une espèce emblématique.

Un vaste ensemble de données a été rassemblé, regroupant 50 variables environnementales et anthropiques, y compris la gestion des forêts, et 424 observations de lucane cerf-volant. Les données sur le coléoptère visaient 242 km<sup>2</sup> sur un total de 3 517 km<sup>2</sup> (6,89%). Les données proviennent de bases de données en ligne, de collections privées et de programmes de sciences citoyennes mis en œuvre sur les territoires français et allemands. Ils couvrent la période 1990-2018, à l'exception des données sur la continuité forestière qui reflète 200 ans de présence forestière.

L'analyse était basée sur trois types de modèles: forêt aléatoire (RF), modèle linéaire généralisé (GLM), régression binomiale négative à zéro inflation (ZI), le dernier prenant en compte la présence de faux zéro. Le facteur qui explique le mieux la distribution du lucane cerf-volant est ANTH, la présence de bâtiments. Cette variable est 100 fois plus influente que toute autre variable. Ceci suggère un biais sérieux dans les données d'observation. Cependant, ce type de biais est régulièrement présent dans toutes les études utilisant la science citoyenne. En conséquence, ANTH doit être considéré comme un substitut de la densité d'observateurs. Ainsi, la carte d'observation reflétait les contacts entre les espèces et les observateurs et non la présence réelle du lucane cerf-volant.

Les modèles et les scénarios proposés ont montré qu'il est possible d'évaluer quels environnements sont favorables au lucane cerf-volant. De plus, d'importantes applications sont possibles, bien que nos tests doivent être considérés comme des exemples de procédures d'évaluation. Ces applications permettent des enquêtes détaillées pour divers scénarios de gestion conçus pour favoriser la présence du lucane cerf-volant. Les modèles créés dans le cadre de cette étude peuvent être utilisés avec n'importe quel jeu de données de distribution, pas seulement le lucane cerf-volant. Ce sont des outils puissants pour lier les observations des espèces aux facteurs environnementaux, pour finalement prévoir la distribution des espèces.

Les études ultérieures sur le lucane cerf-volant devraient porter sur un échantillonnage plus complet. Il a été montré que les relevés n'avaient pas suffisamment couvert certains habitats et n'étaient pas homogènes du point de vue spatial. Par conséquent, trop de faux négatifs (absence d'observations) sont présents dans les données d'observation. Idéalement, de nouvelles observations devront être faites aléatoirement, sur tous les types d'utilisation et de statut des terres (zones protégées, forêt ancienne, etc.). Le taux d'observation global devrait être augmenté pour couvrir environ 457 km<sup>2</sup> supplémentaires si l'on voulait atteindre la pression d'observation obtenue dans une autre étude conduite en Belgique (20,5%, Thomaes *et al.* 2008). Cet objectif peut être difficile à atteindre, sauf en se concentrant sur de petites zones sur lesquelles des recherches approfondies doivent être menées. Les modèles peuvent donc être calibrés sur ces petites zones puis appliqués à l'ensemble de données et au paysage. Les nouvelles enquêtes pourraient tirer parti du réseau d'observateurs mis en place avant l'étude et mis en avant au cours de son développement. Les petites zones doivent être choisies loin des villes (> 2 km) et inclure des zones protégées et des contextes paysagers variés. Les animations événementielles et une communication sur les programmes existants d'observation du lucane cerf-volant menés tant du côté français que du côté allemand de la réserve de biosphère transfrontalière doivent être poursuivies.

## Abstrakt

Diese Studie betrachtet das Gebiet des grenzüberschreitenden Biosphärenreservats Nordvogesen - Pfälzerwald. Allgemeines Ziel ist es, festzustellen, ob dieser Raum die notwendigen Bedingungen für das Bestehen saproxylierter Arten und insbesondere des Hirschkäfers, einer emblematischen Art, gewährleistet.

Es wurde ein großer Datensatz zusammengestellt, der 50 Umwelt- und anthropische Variablen, einschließlich der Waldbewirtschaftung, und 424 Beobachtungen von Hirschkäfern betrachtet. Die Hirschkäferdaten wurden auf 242 km<sup>2</sup> von insgesamt 3 517 km<sup>2</sup> (6,89%) gesammelt. Die Daten stammen aus Online-Datenbanken, privaten Sammlungen und Citizen Science Projekten aus französischen und deutschen Gebieten. Sie gelten für den Zeitraum 1990-2018, mit Ausnahme der Waldkontinuität, die 200 Jahre Waldbestand widerspiegelt.

Die Analyse basierte auf drei Modelltypen: Random Forest (RF), Generalized Linear Model (GLM), Zero-Inflated Negative Binomial Regression (ZI), wobei die letzte das Vorhandensein von False-Null berücksichtigt. Der Faktor, der die Verteilung des Hirschkäfers am besten erklärt, war ANTH, die Anwesenheit von Gebäuden. Diese Variable ist 100 mal einflussreicher als jede andere Variable. Dies deutet auf eine schwerwiegende Verzerrung der Beobachtungsdaten hin. Diese Art von Verzerrung ist jedoch in allen Citizen Science Studien regelmäßig vorhanden. Infolgedessen muss ANTH als Ersatz für die Dichte des Beobachters angesehen werden. So spiegelte die Beobachtungskarte die Kontakte zwischen Arten und Beobachtern wider und nicht die tatsächliche Anwesenheit des Hirschkäfers.

Die vorgeschlagenen Modelle und Szenarien haben gezeigt, dass es möglich ist, zu beurteilen, welche Bedingungen für den Hirschkäfer günstig sind. Darüber hinaus sind wichtige Anwendungen möglich, wobei unsere Tests als Beispiele für Bewertungsverfahren zu betrachten sind. Diese Anwendungen ermöglichen detaillierte Untersuchungen für verschiedene Managementszenarien, die so konzipiert sind, dass sie die Anwesenheit des Hirschkäfers begünstigen. Darüber hinaus können die Modelle mit jedem beliebigen Verteilungsdatensatz verwendet werden, nicht nur mit dem Hirschkäfer. Sie sind starke Werkzeuge, um Artenbeobachtungen mit Umweltfaktoren zu verknüpfen und schließlich die Verteilung der Arten vorherzusagen.

Weitere Studien zum Hirschkäfer sollten sich auf eine vollständigere Probenahme konzentrieren. Es zeigte sich, dass die Erhebungen bestimmte Lebensräume nicht ausreichend abdeckten und nicht räumlich homogen waren. Folglich sind zu viele falsche Negative (Fehlen von Beobachtungen) in den Beobachtungsdaten vorhanden. Im Idealfall werden neue Beobachtungen zu allen Arten der Landnutzung und des Flächenzustands (Schutzgebiete, Altwälder usw.) nach dem Zufallsprinzip durchgeführt. Die Gesamtbeobachtungsrate sollte auf etwa 457 neue km<sup>2</sup> angehoben werden, wenn man den in Belgien erzielten Beobachtungsdruck erreichen will (20,5 %, Thomaes et al. 2008). Dieses Ziel könnte schwer zu erreichen sein, außer durch die Konzentration auf kleine Bereiche, die intensiv untersucht werden sollen. Modelle können daher auf diesen kleinen Flächen kalibriert und dann auf den gesamten Datensatz und die Landschaft angewendet werden. Neue Untersuchungen können von dem vor der Studie eingerichteten und während ihrer Entwicklung geförderten Beobachternetz profitieren. Kleine Gebiete sollten fernab von Städten (> 2 km) gewählt werden und Schutzgebiete und vielfältige Landschaftsstrukturen umfassen. Dies ermöglicht auch die Durchführung ereignisorientierter Animation und Kommunikation über das Monitoringprogramm zum Hirschkäfer auf französischer und deutscher Seite des Biosphärenreservates.

## Lexique et acronymes

**GAM** : Generalised Additive Model.

**GLM** : Generalised Linear Model.

Les **îlots de vieillissement** sont des portions de forêt où, pendant une certaine période, les âges ou les diamètres d'exploitabilité sont augmentés (par exemple doublés).

Les **îlots de sénescence** sont des portions de forêt (0,5 et 3 ha), qui ne seront plus exploitées durant au moins 30 ans. Ils sont à la fois des réservoirs de biodiversité mais aussi, en étant connectés, des corridors écologiques intra-forestiers. Les bois de feuillus doivent représenter au moins 10 arbres/ha dans l'îlot avec des diamètres à 1,20 m de plus de 50 cm ou avoir des éléments importants pour la biodiversité comme une ou plusieurs cavités, des blessures au tronc, des parties cariées, un décollement d'écorce, de grosses branches mortes, etc. La perte financière liée à la non-exploitation de ces surfaces est prise en compte par le versement au propriétaire d'une compensation financière (projet LIFE Biocorridors mai 2016)

**OPIE** : Office pour les Insectes et leur Environnement.

**RBT** : Réserve de Biosphère transfrontalière.

**Saproxyle** (de *sapros* pourri et *xylo* bois). « *Espèces impliquées dans ou dépendantes du processus de décomposition fongique du bois, ou des produits de cette décomposition, et qui sont associées à des arbres tant vivants que morts. Par convention deux autres regroupements d'organismes sont inclus dans cette définition : i) les espèces associées aux écoulements de sève et à leurs produits de décomposition, et ii) les organismes autres que les champignons qui se nourrissent directement du bois.* » (Alexander, 2008).

# 1. Rappels de contexte et d'objectifs

Le Parc Naturel Régional des Vosges du Nord et le Biosphärenreservat Pfälzerwald ont été reconnus par l'UNESCO en 1998 « Réserve de Biosphère transfrontalière des Vosges du Nord - Pfälzerwald ». Cet espace couvre 3 105 km<sup>2</sup>, 1 800 km<sup>2</sup> côté allemand et 1 300 km<sup>2</sup> côté français. Une majorité de cette surface est forestière, principalement de la forêt publique.

La coopération transfrontalière entre les deux Parcs existe depuis de nombreuses années au travers de projets communs dont le programme « LIFE Biocorridors » de février 2016 à mai 2020.

Cette étude porte sur les actions du LIFE Biocorridors en faveur des îlots de sénescence. Le thème central est l'analyse des relations entre les îlots, et plus globalement les caractéristiques de l'environnement et des organismes clés des boisements anciens et vieillissants : les coléoptères saproxyliques dont le Lucane cerf-volant a été choisi comme espèce porte drapeau.

Un travail liminaire d'audit du réseau d'îlots de sénescence coté français a été mené entre 2015 et 2017 (Demets & Bruciamacchie 2017). Il intègre deux scénarios pour placer de nouveaux îlots de sénescence, sur la base de deux modes d'analyse poussée des boisements existants.

La première phase de l'étude (Dodelin *et al.* 2018), a permis de rassembler une partie des données environnementales, la plupart des données de distribution des espèces cibles et a proposé un premier test de modélisation (GLM) à l'échelle des deux espaces protégés. Cette première phase a servi de matière pour le mémoire de Master Sciences de l'Environnement Terrestre 1<sup>ère</sup> année de R. Saurat (2018).

Afin de ne pas surcharger le présent document, nous renvoyons au rapport d'étude précédent pour plus de détails sur les données et les traitements analytiques.

Dans le prolongement de ce premier rendu, les échanges avec les différents acteurs de ce dossier ont permis de rassembler un plus grand nombre de données environnementales pertinentes afin d'améliorer la modélisation : types forestiers, îlots de forêts protégées et climat. Nous avons également profité de cette nouvelle étape de travail pour mettre à jour les données de Lucane cerf-volant avec les observations effectuées sur la seconde moitié de l'année 2018 (enquête OPIE<sup>1</sup> et Artenfinder<sup>2</sup>).

Sur un plan plus technique, deux nouveaux modèles ont été adjoints au "Generalized Linear Model" : "Random Forest" ainsi qu'un modèle de type "Zero Inflated". L'analyse elle-même se découpe en grandes étapes qui ont vu tout d'abord la recherche des variables environnementales qui expliquent le mieux les observations de Lucane cerf-volant. Celles-ci ont permis, dans un second temps, de calculer une prédiction de présence du Lucane sur tout le territoire. Certaines de ces variables constituent également des leviers sur lesquels il est possible d'agir ou de développer une politique générale en faveur des Lucanes. Cette discussion fait l'objet de la dernière phase d'analyse de nature prospective, durant laquelle des scénarios d'aménagement sont envisagés et discutés sur la base des prédictions de présence du Lucane, et de leur confrontation avec le terrain.

---

1 [www.insectes.org/enquete/Lucane-cerf-volant.html](http://www.insectes.org/enquete/Lucane-cerf-volant.html)

[www.facebook.com/parcvosgesnord/photos/a.276322475752393.93872.276312969086677/1502810763103552/?type=3](https://www.facebook.com/parcvosgesnord/photos/a.276322475752393.93872.276312969086677/1502810763103552/?type=3)

2 [www.artenanalyse.net/artenanalyse/](http://www.artenanalyse.net/artenanalyse/)



## 2. Méthodes d'analyses

### a) Démarche globale de l'analyse

La base scientifique suivie dans cette étude, est le travail réalisé en Belgique par A. Thomaes et son équipe (Thomaes 2008, 2009, Thomaes *et al.* 2008, 2010). Ces travaux ont servi de référence pour le choix des modèles statistiques et des grand types de variables environnementales. Ils ont été présentés en détail dans le rapport d'étude précédent. Les variables pertinentes issues de ces travaux sont données dans le Tableau 1.

Tableau 1. Variables environnementales ayant un impact significatif sur la distribution du *Lucane cerf-volant* en Belgique selon Thomaes *et al.* (2008)

Variable	Relation avec le <i>Lucane cerf-volant</i>
Surface agricole	- - -
Surface urbanisée	+ + +
Altitude	+ +
Sol non ennoyé au cours de l'année	+ +
Surface de forêts de feuillus ou mixtes	+ +
Surface de conifères	~ +
Température moyenne annuelle	~ +
Hauteur de pluie annuelle	~ +
Surface de fruticées	~ -

Notre analyse consiste à croiser les variables environnementales et les données de distribution du *Lucane cerf-volant*. Ces données sont présentées en détail au prochain chapitre.

Dans un premier temps les variables environnementales sont testées à la recherche de corrélations entre elles. Le but de cette première étape est d'éliminer les variables redondantes pour alléger les calculs suivants et renforcer les interactions entre les variables restantes.

La seconde étape est effectuée à l'aide de modèles qui recherchent les liens entre la distribution du *Lucane* et les variables environnementales. De nouveau, cette étape va retirer les variables dont le pouvoir explicatif n'est pas significatif pour aboutir à une sélection d'un petit groupe de variables explicatives de la distribution observée.

À partir de ce groupe restreint de variables, les modèles permettent le calcul de prédictions de présence du *Lucane*, simplement en prenant en compte l'état de l'environnement. Ces prédictions peuvent ensuite être confrontées à des données d'intérêt et documenter des scénarios d'aménagement en faveur du *Lucane*. Nous chercherons par exemple à savoir si les prédictions obtenues sont en relation avec le réseau d'îlots de sénescence existant. Si ce réseau était augmenté, cette relation serait-elle améliorée ? Etc.

### b) Données

#### (1) Structuration des variables

Toutes les données sont rapportées à un maillage kilométrique qui sert de référentiel unique. Le système Européen EPSG:3035 a été retenu (European Environment Agency), avec pour l'espace d'étude, 3 517 cellules de 1 km<sup>2</sup>.

Les variables sont toutes exprimées à l'échelle de la maille kilométrique. Chaque maille dispose ainsi d'une

"carte d'identité environnementale multi-critères", à laquelle sont ajoutées les informations de présence / pseudo-absence des coléoptères. Nous avons considéré les absences d'observations non pas comme des absences vraies mais comme des pseudo-absences car beaucoup de ces zéro découlent d'absences d'observations.

Les données cartographiques ont été traitées avec le logiciel QGIS 2.18.16.

## (2) Variables environnementales

Le Tableau 2 résume les variables environnementales retenues pour l'étude ainsi que leurs sources. Nous donnons ci-dessous quelques détails importants à propos de ces variables.

Toutes les variables s'inscrivent dans la période actuelle (<20 ans). Seule la continuité historique de présence forestière recouvre une période d'environ 200 ans. Pour définir cette continuité en France, et compte tenu des plages temporelles disponibles pour l'Allemagne, la carte de Cassini a été délaissée au profit de la carte d'État Major, plus solide géographiquement.

Les données de topographie sont pour certaines, des indices exclusifs développés par Geopeka et dont on trouvera plus de détails à l'adresse <https://notes.mapkiwiz.fr/bdaura/metrics/index.html>.

Les surfaces d'espaces protégés ont été rapportées à un espace de 8 mailles réparties autour de chaque maille, soit un rayon de 1,5 km autour du centre de chaque maille. Ce rayon de 1,5 km correspond à la distance moyenne de dispersion des Lucanes cerf-volants (Illustration 1). En quelque sorte, les Lucanes présents dans ce buffer peuvent potentiellement interagir avec la surface protégée correspondante.

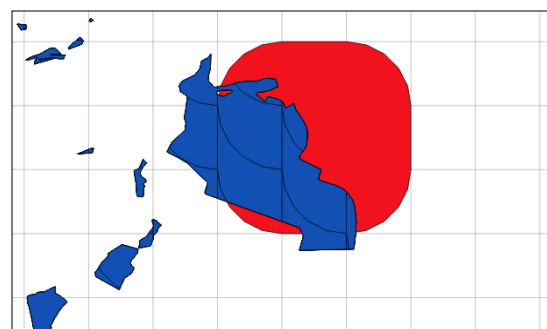


Illustration 1. Buffer (en rouge) utilisé pour le calcul des surfaces d'aires protégées autour de chaque maille

De nombreux jeux de données ont été collectés mais écartés pour l'analyse. Les principales raisons de ce rejet tiennent dans le manque de précision ou, plus souvent, dans l'absence de données équivalentes dans l'autre pays. C'est par exemple le cas du bois mort qui a fait l'objet d'inventaires en territoire français mais pour lequel il n'existe pas de données équivalentes en Allemagne. Enfin, le réseau d'îlots de sénescence n'a pas été rendu disponible pour le territoire allemand.

Tableau 2. Données environnementales retenues pour l'analyse

Type de donnée	Abréviation	Descriptif	Source
Occupation du sol	ANTH	Espaces urbanisés (anthropiques)	Corine Land Cover (2012) <a href="https://land.copernicus.eu/pan-european/corine-land-cover">https://land.copernicus.eu/pan-european/corine-land-cover</a>
	CULT	Grandes cultures	
	RG	Cultures en rang (vignes, vergers)	
	FF	Forêt de feuillus	
	FC	Forêt de conifères	
	FM	Forêt mixte	
	TRANS	Transition arbustif vers arboré	
	PAIRIES	Prairie + Prairie naturelle	
	TOUR	Tourbière - bruyère	
	EAU	Zones en eau	
Indice Culture Prairie	ICP2	$(CULT+RG)/(P+PN)$	
Indice de Fermeture du Paysage	IFP_2	$(FF+FC+FM)/[TOTAL-(EAU+ANTH)]$	
Linéaire de lisière	LISIEREL	Fusion des polygones FF, FC, FM, puis calcul du	

Type de donnée	Abréviation	Descriptif	Source
		linéaire par maille	
Capacité en eau du sol	AWC_TOP AWC_SUB	Capacité en eau du sol de surface Capacité en eau du sous-sol	
Pédologie de surface (0-30 cm)	TOP_CLAY TOP_SILT TOP_SAND TOP_GRAVEL	Proportion d'argile Proportion de limon Proportion de sable Proportion de gravier	European Soil Data Centre <a href="https://esdac.jrc.ec.europa.eu/">https://esdac.jrc.ec.europa.eu/</a>
Espaces protégés	Aprotect	Surfaces protégées dans les 8 mailles autour de chaque maille (cf. texte)	
		Espaces pris en compte pour le calcul : Îlots de sénescence Îlots de vieillissement Arrêtés préfectoraux de protection de biotope (APPB) Réserves Biologiques Forestières (Intégrales et Dirigées) Réserves naturelles nationales (RNN) Espaces Naturels Sensibles (ENS) Kernzone	PNR des Vosges du Nord Biosphärenreservat Pfälzerwald
Boisement de chêne	Achene	Surfaces de chêne dominant : Chênes Décidus Lichte Eichen (chêne vieux en sous-bois clair), qui intègre une variante Buche Unterstand (chêne avec sous-bois de hêtre)	Demets & Bruciamacchie (2017) <a href="http://landesforsten-rheinland-pfalz.de">landesforsten-rheinland-pfalz</a> <a href="http://wald-rlp.de">http://wald-rlp.de</a>
Ancienneté des forêts	AforetAnc	% de forêt présente depuis 200 ans (brèves coupes à blanc possibles ; résolution maximale ~50 ha). % de forêt de création récente, durant le 19e ou le 20e siècle (résolution maximale ~50 ha). Cartes exploitées : <ul style="list-style-type: none"> <li>Histwald (Historisch alte Waldstandorte)</li> <li>Carte d'État Major - milieu XIXe</li> <li>Carte topographique - 1950/1960</li> <li>Carte topographique - actuel</li> </ul>	Bundesamt für Naturschutz <a href="http://www.bfn.de">www.bfn.de</a> (Glaser & Hauke 2004) Institut Géographique National (IGN)
Relief	alt_moy alt_min alt_max pente expo profil tpi_1k	Altitude moyenne Altitude minimum Altitude maximum Indice de pente en pourcentage, dérivée de l'altitude. Indice d'exposition. Angle de la projection horizontale de la normale au MNT avec l'axe nord du système de référence spatiale, exprimé en degrés décimaux : 0° = plein nord, 180° = plein sud, etc. Indice de profil. Dérivée seconde de l'altitude (courbure). Une valeur fortement positive indique une position en sommet, une valeur fortement négative une position dans un creux, les valeurs autour de 0 un plat ou une pente uniforme Indice de position topographique, résolution 1 km (micro-relief). Une valeur fortement positive indique une position en sommet, une valeur fortement négative une position dans un creux, les valeurs autour de 0 un plat ou une pente uniforme à l'échelle de calcul de l'indicateur.	IGN : BD ALTI® 50 m <a href="http://professionnels.ign.fr/bdalti">http://professionnels.ign.fr/bdalti</a>

Type de donnée	Abréviation	Descriptif	Source
	tpi_2k	Indice de position topographique, résolution 2 km (macro-relief)	
	Ifrm_moy	Valeur moyenne dans la maille de la catégorie d'indice de forme topographique. Les catégories attribuées aux types de relief sont :  0 → canyon, vallée en forme de V 1 → pente latérale en situation de versant 2 → incision/drain en situation de sommet 3 → vallée en forme de U 4 → plaine large 5 → versant de pente uniforme 6 → sommet aplati 7 → surélévation locale d'un fond de vallée 8 → ligne de crête en situation de versant 9 → crête étroite	
Climat	bio1mean	T° moyenne annuelle	
	bio2mean	Moyenne des écarts diurnes de T° (moyenne des maximales et minimales mensuelles)	
	bio4mean	Saisonnalité de T° (% de déviation standard)	
	bio5mean	T° maximale du mois le plus chaud	
	bio6mean	T° minimale du mois le plus froid	
	bio7mean	Gamme de T° annuelle (BIO5-BIO6)	
	bio8mean	T° moyenne du trimestre le plus humide	
	bio9mean	T° moyenne du trimestre le plus sec	
	bio10mean	T° moyenne du trimestre le plus chaud	WorldClim - Global Climate Data
	bio11mean	T° moyenne du trimestre le plus froid	www.worldclim.org/bioclim
	bio12mean	Précipitation annuelle	
	bio13mean	Précipitation du mois le plus humide	
	bio14mean	Précipitation du mois le plus sec	
	bio15mean	Saisonnalité des précipitations (coef. de variation)	
	bio16mean	Précipitation du trimestre le plus humide	
	bio17mean	Précipitation du trimestre le plus sec	
	bio18mean	Précipitation du trimestre le plus chaud	
	bio19mean	Précipitation du trimestre le plus froid	
Pays	Pays		

### **(3) Mise à jour des données de Lucane cerf-volant**

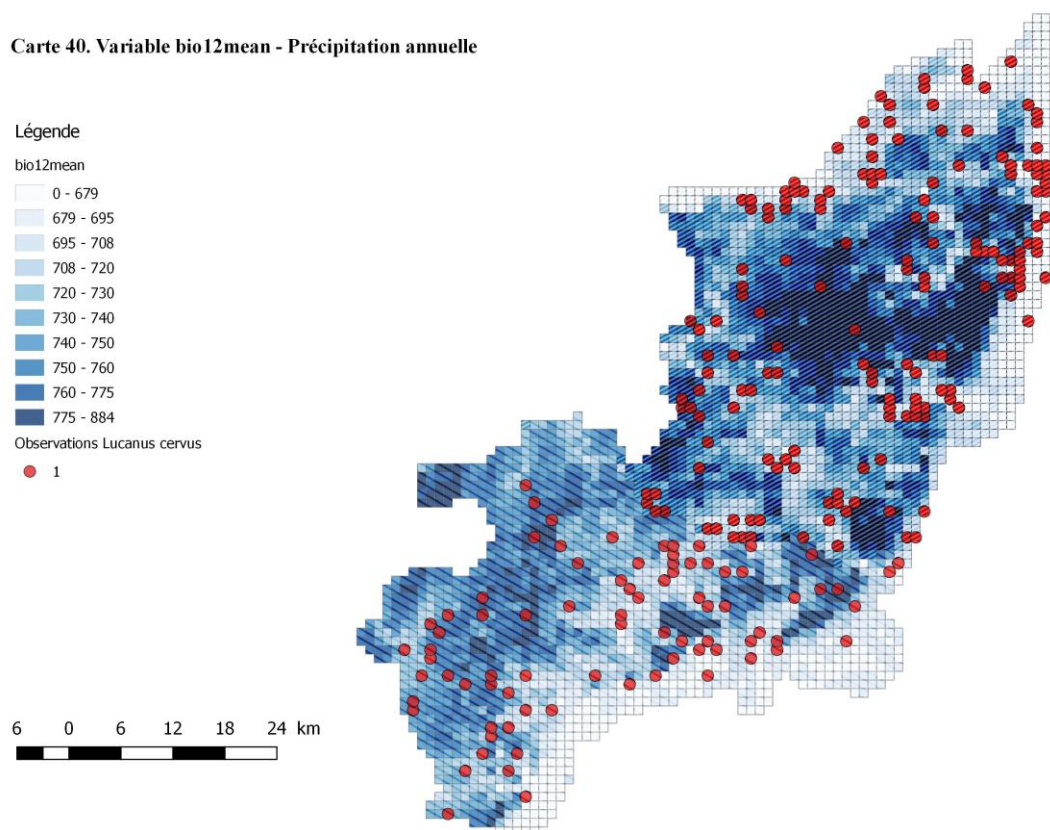
Les données de Lucane cerf-volant ont été complétées pour l'année 2018 sur la base de deux jeux de données disponibles et eux-mêmes à jour : L'enquête Lucane de l'OPIE<sup>3</sup> dans le PNR des Vosges du Nord pour l'année 2018 et le site [www.artenfinder.de](http://www.artenfinder.de).

Les données bancarisées ont été validées chaque fois que cela fut possible (photographies associées, fiabilité des auteurs ou des validateurs des observations). Par exemple, deux données proposées par [artenfinder](http://www.artenfinder.de) ont été invalidées car les photographies associées montraient *Dorcus parallelipedus*. Cela a été signalé au gestionnaire des données en ligne. Les données anciennes de plus de 20 ans ont été écartées pour ne pas introduire de trop grand décalage avec les observations et les descripteurs de l'environnement, qui procèdent de données récentes. Dans le jeu de données Lucane, une dizaine de données datent de 1990 à 2005, une majorité se place entre 2005 et 2018.

3 [www.insectes.org/enquete/Lucane-cerf-volant.html](http://www.insectes.org/enquete/Lucane-cerf-volant.html)

En novembre 2018, le nombre d'observations d'individus de Lucane pour les analyses se porte à 424, soit une présence sur 242 mailles. Les 424 observations de Lucane cerf-volant ne peuvent pas être toutes retenues pour la modélisation. Une petite moitié, 182 observations, correspond à une information redondante : soit d'un même Lucane revu plusieurs fois de suite au même endroit, soit d'un même observateur réalisant d'une année sur l'autre, ses observations sur des lieux où l'on sait que le Lucane est présent, soit enfin que notre regroupement des données par maille rattache à une même maille des observations proches. Ces éléments créent un effet d'amplification qui exacerbe les résultats des modélisations et biaise l'analyse en faveur des caractéristiques portées par mailles les plus riches en observations.

Par exemple, les trois mailles regroupant le plus d'observations concentrent 24, 20 et 13 observations alors que 169 mailles ne comptent qu'une observation. Notre solution a donc été de travailler sur une donnée d'observation "résumée" : la présence / absence par maille, ou pour être plus précis, la présence / pseudo-absence par maille car les absences d'observations ne sont pas, dans la plupart des cas, des absences réelles du Lucane mais des absences d'observateurs. La distribution statistique de ce type de données est de type binomiale.



*Illustration 2. Distribution des 424 observations de Lucane dans l'aire d'étude et précipitations annuelles (BIO12mean). Données de novembre 2018*

#### **(4) Données de coléoptères saproxyliques**

Le terme saproxylique délimite les contours d'un groupe écologique et non taxonomique. Celui-ci renferme une grande diversité d'organismes avec entre 2 500 et 2 900 espèces en France et vraisemblablement en Allemagne (Dodelin 2010). Les coléoptères saproxyliques comptent dans leurs rangs des espèces rares et/ou très exigeantes, d'autres sont limitées à des milieux préservés de forts impacts humains. Plusieurs catalogues Dodelin, Fantino, Darmon 2018. *Lucanus cervus* dans le PNR des Vosges du Nord - Pflzerwald

cherchent à rassembler ces espèces ayant des propriétés indicatrices de l'état et de la naturalité des boisements (Brustel 2001, Müller *et al.* 2005, Sebek *et al.* 2012, Eckelt *et al.* 2017).

Il était envisagé de compléter les résultats obtenus avec le Lucane cerf-volant, en appliquant l'analyse à un groupe d'espèces saproxyliques indicatrices. Des données ont été obtenues pour 540 espèces mais elles sont très ponctuelles dans le temps et dans l'espace. Il s'agit principalement de données de piégeages caractérisées par un grand nombre d'espèces obtenues sur un très petit nombre de stations. De plus, ces stations sont pour la plupart, situées dans des espaces protégés ce qui exclue tout comparatif avec d'autres types de statuts administratifs.

Les recherches et sollicitations pour l'obtention d'autres jeux de données n'ont pas abouti. Comme cela fut le cas ponctuellement avec certaines données de Lucane cerf-volant, des acteurs ont été réticents pour communiquer leurs jeux de données. Dans le même temps, les deux principaux fournisseurs de données<sup>4</sup> ne proposent qu'une résolution spatiale à la maille de 10\*10 km ou de 5\*5 km. À cela s'ajoute le fait que la plupart des espèces discrètes ou méconnues sont très peu documentées dans ces deux bases de données publiques, alors qu'il s'agit justement des espèces pour lesquelles nous aurions besoin de données. Enfin, beaucoup de données sont très anciennes, largement plus que notre limite de 20 ans, si bien qu'il ne reste presque aucune information adaptée à nos contraintes d'analyses. Sans parler des possibles erreurs liées aux données elles-mêmes dans ce type de base agrégeant un grand nombre de sources (erreurs d'identification et de localités principalement). Un essai de collecte d'information a été tenté à partir de sources bibliographiques mais l'entreprise est colossale pour pouvoir couvrir l'ensemble des deux Parcs, avec une information géographique souvent insuffisante (à la commune ou à la maille) voire ingérable (synonymie, doublons de lieux-dits, erreurs de liens entre lieux-dits et communes, etc.).

En conclusion de cette enquête de faisabilité de l'usage de données de coléoptères saproxyliques, nous sommes contraints d'abandonner cette partie du projet.

### *c) Choix des modèles*

Les deux modèles les plus performants d'après Thomaes *et al.* (2008) étaient "**Random Forest**" (**RF**), un modèle développé récemment qui utilise l'apprentissage machine via des réseaux de neurones artificiels (Cutler *et al.* 2012), puis venait un modèle plus éprouvé, le "**Generalized Linear Model**" (**GLM**).

Un troisième modèle a été utilisé pour tenir compte des observations biaisées dans le temps et dans l'espace (toute la carte n'ayant pu être suivie tout le temps), avec comme conséquence de nombreux espaces vides de données (zéros). Ces fausses absences d'observation sont correctement traitées par un modèle de la famille des "Generalized Additive Model" (**GAM**), le "**Zero-Inflated Negative Binomial regression**" (**ZI**). Une telle approche a donné de bons résultats pour l'analyse à grande échelle (le Portugal) de données d'observations ponctuelles obtenues par des réseaux de naturalistes amateurs (Tiago *et al.* 2017).

Ces modèles RF, GLM et ZI relient les observations de Lucane cerf-volant à la somme des variables selon une formule du type :

$$\text{Observations} = f(\text{var.1} + \text{var.2} + \dots)$$

Les analyses ont été réalisées au moyen du logiciel R vers.3.4.41, RStudio vers.1.1.442 et de PAST vers.3.18 (Hammer 2017).

#### **(1)RF - Random Forest**

RF est un modèle puissant et adapté aux grands jeux de données. Son objet est de rendre compte des variables ayant une importance pour expliquer les observations. Il classe hiérarchiquement les variables entre-elles et construit donc des arborescences, répétées aléatoirement un grand nombre de fois (d'où le nom donné au

4 [www.inpn.mnhn.fr](http://www.inpn.mnhn.fr) ; [www.colkat.de](http://www.colkat.de)

modèle, voir James *et al.* 2013 pour plus de détails techniques). RF calcule ainsi la participation de chaque variable à la construction des arborescences et établit des écarts entre variables, écarts qui vont permettre d'écarter les variables les moins importantes. Les deux indices qui correspondent à l'apport explicatif de chacune dans le modèle, sont %IncMSE et IncNodePurity.

Le premier modèle est calculé avec les données retenues d'après les résultats des recherches de corrélations.

Cette première analyse écarte les variables dont les valeurs de %IncMSE sont négatives. De telles valeurs indiquent que lorsque les valeurs de la variable sont mélangées (permutées aléatoirement), elles participent mieux au modèle que dans leurs versions non-permutées, ce qui est un non-sens pour notre étude.

Le second modèle est donc construit en retirant les variables ayant des valeurs de %IncMSE négatives. À l'issue de cette classification, nous proposons de ne conserver que les variables qui participent le plus au modèle afin d'améliorer les performances de celui-ci. Nous définissons comme limite inférieure la moyenne des valeurs de IncNodePurity : les variables ayant des valeurs supérieures sont seules conservées.

Le troisième modèle, final, est construit à partir des variables restantes à l'issue des deux sélections.

### **(2)GLM - Generalized Linear Model**

Comme dans le cas de RF, le GLM est construit à partir des variables retenues à la suite de la phase de recherche de corrélations.

La sélection des variables pertinentes se déroule en une étape. Elle s'appuie sur la valeur du Critère d'Information de Akaike (AIC). Le modèle complet produit un AIC initial. Puis le modèle est de nouveau construit en éliminant une variable (commande `stepAIC()` de R). La nouvelle valeur de l'AIC est comparée à l'initiale : si l'AIC baisse, alors la variable n'est pas pertinente et peut être retirée, si l'AIC augmente alors la variable doit être conservée. Selon le principe de parcimonie, n'est retenu que le modèle possédant un AIC maximal, tout en faisant appel à un minimum de variables. Ce modèle final intègre ainsi les variables ayant les plus fortes capacités explicatives.

### **(3)ZI - Zero-Inflated Negative Binomial regression**

Les premières approches analytiques ont montré que nos données d'observations de Lucane souffrent d'un très fort "biais observateur". Ce biais se transcrit par un plus grand nombre de contacts entre l'Homme et le Lucane à proximité des villes et villages tandis que les cœurs de massifs forestier n'ont pas - ou très peu - été prospectés. Un tel biais se retrouve dans presque toutes les études de type "science participative" et il est très difficile de s'en défaire (Tyre *et al.* 2003, Tiago *et al.* 2017). Du point de vue de l'analyse, cela se traduit, dans les données d'observations de Lucane, par des zéros en "surnombre", c'est à dire que beaucoup d'entre-eux ne reflètent pas une réelle absence de l'espèce comptée, mais plus certainement l'absence d'observateurs (zéros comptés contre faux zéros).

Le modèle ZI fonctionne globalement comme une double GLM : la première étape prend en compte les variables sans modifications (les zéros sont considérés comme comptés), la seconde étape intègre la présence d'un grand nombre de faux zéros au moyen d'un tirage aléatoire qui permet d'estimer leur importance dans les données (Zeileis *et al.* 2008). Dans notre cas, nous suivrons attentivement la variable ANTH que l'on peut considérer comme un reflet de la densité des observateurs potentiels. Les variables significatives lors de la première étape mais qui ne le sont plus dans la seconde peuvent être retirées du modèle en considérant que l'impact des faux zéro est trop important.

### **(4)Prédictions de distributions**

Chaque modèle ayant permis de relier en une équation les variables environnementales avec la distribution du Lucane, il est possible en retour d'utiliser cette équation pour déterminer les mailles où les facteurs environnementaux prennent des valeurs favorables à la présence du Lucane (commande `predict()` de R).

Les trois modèles n'ayant pas exactement sélectionné les mêmes variables, ni même produit les mêmes

coefficients dans leurs équations finales respectives, les prédictions présentent nécessairement des différences. À la manière de Thomaes *et al.* (2008), nous avons synthétiser les trois prédictions en une prédiction finale. Techniquement, ce regroupement fait appel à une sélection des prédictions "acceptables" dans notre contexte. Un choix de seuil doit être fait à ce niveau pour écarter ou conserver les prédictions. Quelques méthodes permettent de fixer statistiquement ce seuil (Perez-Jaume *et al.* 2017), méthodes qui nous sont inaccessibles au vu de nos jeu de données et de la complexité de l'opération.

Nous reprenons donc ici la technique utilisée par Thomaes *et al.* (2008), qui exploite la fréquence d'occurrence de l'espèce cible (prevalence criteron ou species' prevalence selon auteurs, cf. Manel *et al.* (1999)). Cette méthode correspond à une version simplifiée des méthodes plus poussées, lesquelles prennent en compte les statistiques liées aux résultats de prédiction (vrais et faux positifs contre vrai et faux négatifs, cf. Manel *et al.* (2002)). Ici on utilise la situation mesurée comme un simulacre des prédictions. Le seuil est donc fixé par :

$$\text{Seuil} = \text{Nombre de mailles avec observations} / \text{Nombre total de mailles}$$

Dans l'étude de Thomaes *et al.* (2008), le seuil est placé à  $73/356 = 0,205$ . Dans notre cas le seuil devient  $242/3517 = 0,0688$ . Cette façon de faire n'est pas tout à fait satisfaisante puisque les grands jeux de données avec peu d'observations donneront des prédictions plus fragiles mais pour lesquelles on conservera les prédictions à des niveaux plus faibles. Il y a donc un risque plus important de conserver une fausse prédiction.

En dernier lieu, les prédictions validées sont les mailles prédites par les trois modèles avec des valeurs dépassant le seuil de 0,0688.

#### ***d) Construction de scénarios***

Une fois les données de prédictions construites, différentes questions peuvent être posées et résolues statistiquement. Par exemple pour savoir si le réseau actuel d'îlots de sénescence du PNRVN recouvre les espaces où le Lucane est prédit. Ou encore pour déterminer si une augmentation de quelques pourcents de la surface protégée permettrait d'inclure significativement plus d'habitats de Lucane qu'actuellement, et ce, en fonction de différentes configurations spatiales.

Pour cela, nous avons créé artificiellement de nouvelles variables correspondant à trois augmentations : +5 %, +10 % et +20 % de l'existant. Ces surfaces supplémentaires sont ajoutées à l'existant selon plusieurs modes opératoires. Ceux ci-ont intervenir d'une part l'ajout de "quelques grandes" ou "beaucoup de petites" entités spatiales, et d'autre part, le caractère aléatoire de l'ajout pour toute la carte ou bien son conditionnement par un réseau de corridors reliant l'existant (cf. les Illustrations 3 à 6) :

- Scénario 1 : dans et autour de l'existant, sans création de nouvelle entité s'il n'en existe pas déjà une ;
- Scénario 2 : ajout de 10 grandes entités de façon régulière sur l'ensemble de la carte, en cohérence avec les grande entités existantes ;
- Scénario 3 : ajout de 10 grandes entités de façon aléatoire au sein de corridors de 2 km de large reliant les espaces protégés existants ;
- Scénario 4 : ajout de 200 petites entités de façon aléatoire sur l'ensemble de la carte ;
- Scénario 5 : ajout de 200 petites entités de façon aléatoire au sein de corridors de 2 km de large reliant les espaces protégés existants.

Les variables traitées selon le scénario 1 sont LISIEREL, Aprotect et Achene. Une seule variable a été traitée selon les autres scénarios : Aprotect.

Le nombre d'entités a été choisi pour être cohérent avec les surfaces protégées existantes. Ainsi les petites entités recouvrent chacune au minimum 1,27 ha pour 5 % d'augmentation, soit la taille habituelle d'un îlot, tandis qu'avec 10 grandes entités, chacune d'elle occupe 100 ha au maximum (20 % d'augmentation), soit à peu-près la taille d'une réserve biologique (Tableau 3). Les entités créées sont circulaires.



Tableau 3. Entités ajoutées aux zones protégées existantes pour la création de scénarios

Facteur d'augmentation	Surface créée (ha)	Surface par entité (ha)		Rayon par entité (m)	
		200 petites	10 grandes	200 petites	10 grandes
+5%	254	1,27	25,4	63	284
+10%	508	2,54	50,8	89	402
+20%	1016	5,08	101,6	127	568

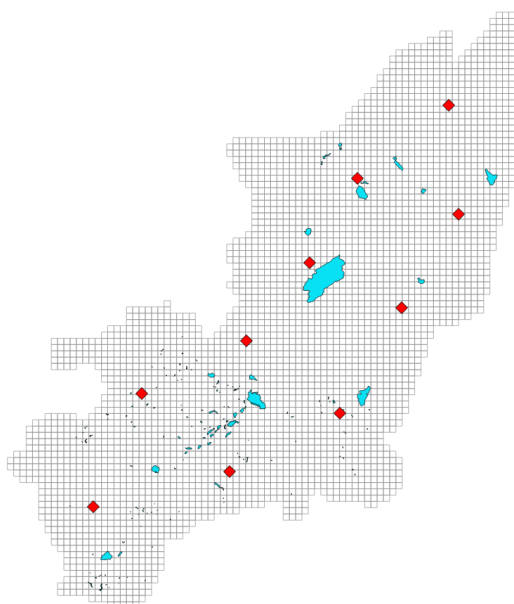


Illustration 3. Schéma de principe du scénario 2

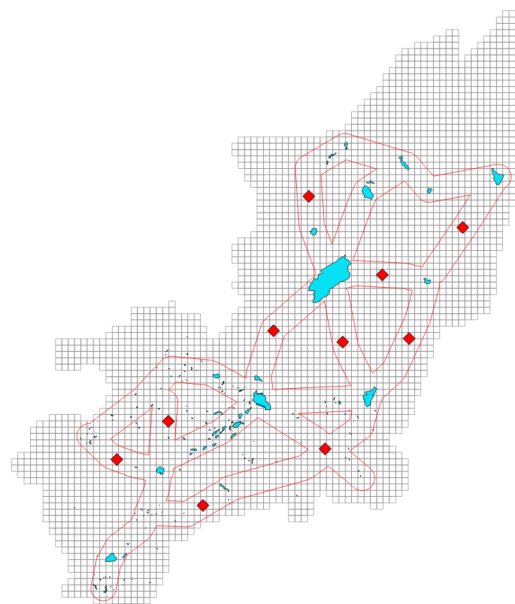


Illustration 4. Schéma de principe du scénario 3

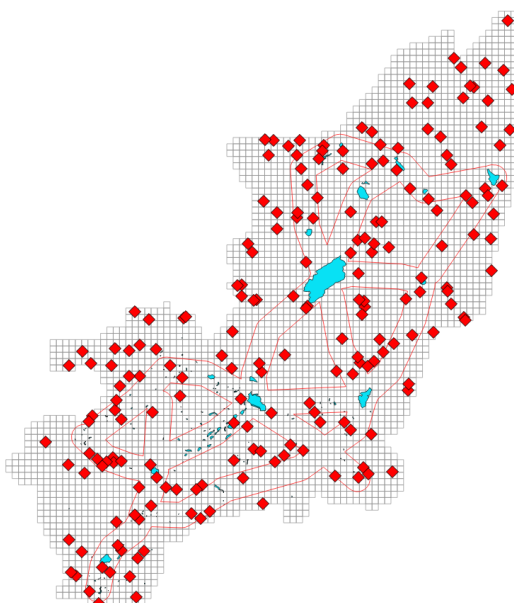


Illustration 5. Schéma de principe du scénario 4

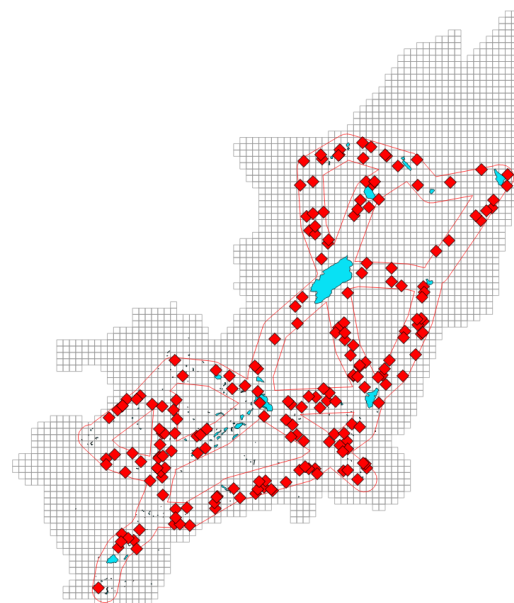


Illustration 6. Schéma de principe du scénario 5

## e) Autres outils statistiques

### (1) Test de Spearman

Le test de corrélation de Spearman ( $r_s$ ) a été utilisé car il permet de prendre en charge tous les types de données, qu'elles suivent ou non la Loi Normale. Ce test non-paramétrique compare les rangs des valeurs. La valeur maximale de  $r_s$  ( $r_s = 1$ ) implique l'identité des deux échantillons.

### (2) Analyse de variance

Certaines des variables utilisées ici, comme les surfaces d'espaces protégés, présentent des valeurs qui ne suivent pas la Loi Normale. Les tests classiques de type Chi2 ou ANOVA ne peuvent s'appliquer. Nous utiliserons à la place une analyse de variance non-paramétrique, c'est à dire qui exploite les classements (rangs) entre valeurs au sein de la variable : le one-way PERMANOVA ou le test de Kruskal-wallis. Ces tests recherchent les différences significatives entre groupes formés au sein d'une variable ou entre deux variables. Par exemple dans notre étude, deux groupes ont été formés pour la variable Aprotect, en distinguant les valeurs prises par Aprotect dans les mailles où le Lucane est prédit (Aprotect groupe 1) et dans les mailles où le Lucane n'est pas prédit (Aprotect groupe 2). Lorsque les données suivent la Loi Normale, le test t de Student ou l'ANOVA ont été employés.

## f) Préparation des données : corrélations entre variables

Le premier travail effectué sur les variables environnementales est une recherche des corrélations au sein des grandes familles de variables (sol, climat, etc.). L'objectif est d'écarter les variables dépendantes ou redondantes et d'alléger le modèle explicatif initial.

### (1) Variables pédologiques

Nous avons retenu les seules données relatives au sol superficiel et écarté les données décrivant le sous-sol. Deux raisons à ce choix : d'une part le caractère incomplet des données du sous-sol, avec de nombreuses mailles non renseignées, et d'autre part, la très forte corrélation entre les mêmes variables issues du sol superficiel (TOP\_) et du sous-sol (SUB\_).

Il existe une corrélation significative et positive entre trois éléments : argile (CLAY), limons (SILT) et cailloutis (GRAVEL) (Tableau 4). Lorsque l'un est présent, les autres le sont aussi. Inversement, lorsque du sable (SAND) est présent, les autres éléments se font plus rares. La corrélation très fortement négative entre argiles et sables permet de retirer la variable TOP\_CLAY du modèle initial.

Tableau 4. Corrélations de Spearman ( $r_s$ ) pour les variables pédologiques  
Seuil de significativité : \*\*\* : 0,001

Variables	TOP_CLAY	TOP_SILT	TOP_SAND
TOP_SILT	0,29***	0	
TOP_SAND	-0,84***	-0,50***	0
TOP_GRAVEL	0,31***	0,47***	-0,42***

### (2) Variables d'occupation du sol

L'indice de cultures et prairies (ICP2) joue très bien son rôle de synthèse pour ces variables. Au sein des plus fortes corrélations (Tableau 5), il est positivement lié à CULT et PRAIRIES, et négativement à l'indice de

fermeture du paysage IFP\_2. Les variables CULT et PRAIRIES peuvent donc être retirées du modèle initial où leurs informations seront portées par ICP2.

La présence d'anciennes forêts (AforetAnc), est significativement corrélée avec plusieurs autres variables : négativement avec les constructions (ANTH) ainsi que les cultures et prairies ; positivement avec les forêts de conifères (FC) et mixtes (FM). Les mêmes corrélations se retrouvent avec les indices synthétiques ICP2 et IFP\_2.

L'information de présence de forêts anciennes étant fondamentale pour notre problématique, il n'est pas possible de la retirer au profit d'une variable plus générale.

*Tableau 5. Corrélations de Spearman de valeurs  $|rs| \geq 0,5$  pour les variables d'occupation du sol  
Les variables n'ayant aucune corrélation dépassant ce seuil ne sont pas présentées  
Seuil de significativité : \*\*\* : 0,001*

Variables	ANTH	CULT	FC	FM	PRAIRIES	ICP2	IFP_2
FC		-0,49***					
ICP2		0,64***			0,66***		
IFP_2		-0,64***	0,49***		-0,56***	-0,51***	
AforetAnc	-0,49***	-0,60***	0,52***	0,50***	-0,52***	-0,49***	0,79***

Pour être complets à propos des corrélations entre les variables d'occupation du sol et les autres familles de variables, il faut signaler quelques corrélations significatives avec la topologie (Tableau 6). Les cultures sont reliées aux altitudes faibles et aux secteurs plats (négativement corrélées à alt\_max et à la pente). Inversement, la pente est corrélée avec les forêts de conifères et les forêts mixtes. Corrélation que l'on retrouve dans l'indice de fermeture du paysage (IFP\_2) et la surface de forêts anciennes.

*Tableau 6. Corrélations de Spearman de valeurs  $|rs| \geq 0,5$  pour les variables de topologie et d'occupation du sol  
Les variables n'ayant aucune corrélation dépassant ce seuil ne sont pas présentées  
Seuil de significativité : \*\*\* : 0,001*

Variables	CULT	FC	FM	IFP_2	AforetAnc
alt_moy				0,50***	0,52***
alt_max	-0,50***			0,62***	0,64***
pente	-0,56***	0,51***	0,53***	0,64***	0,66***

### **(3) Variables de topologie**

Au sein des variables de topologie (Tableau 7), les altitude minimales, moyennes et maximale sont fortement corrélées positivement entre-elles. De plus, alt\_moy et alt\_max sont positivement corrélées avec la pente. Il est ainsi possible de retirer du modèle initial les variables alt\_moy et alt\_max dont l'information sera portée par les variables alt\_min et pente.

Les indices de topographie (profil, tpi\_1k, tpi\_2k et Ifrm\_moy) sont très fortement inter-corrélés en raison de leur mode de construction. Un seul peut être conservé dans le modèle initial. Nous avons retenu l'indice de forme topographique (Ifrm\_moy) qui porte une information originale par rapport aux autres indices construits avec un lien fort vis-à-vis de l'altitude.

Tableau 7. Corrélations de Spearman de valeurs  $|rs| \geq 0,5$  pour les variables de topologie  
 Les variables n'ayant aucune corrélation dépassant ce seuil ne sont pas présentées  
 Seuil de significativité : \*\*\* : 0,001

Variabiles	alt_moy	alt_min	alt_max	profil	tpi_1k	tpi_2k
alt_min	0,87***					
alt_max	0,87***	0,62***				
penste	0,51***		0,77***			
tpi_1k				0,82***		
tpi_2k				0,80***	0,82***	
lfrm_moy	0,59***			0,69***	0,81***	0,71***

#### (4) Variabiles climatiques

Les corrélations dans le groupe des variables climatiques sont nombreuses. La plupart (BIO9mean exceptée), sont également reliées à l'altitude de manière évidente.

Pour les températures (variables BIO1mean à BIO11mean, Tableau 8), la moyenne des écarts diurnes mensuels de température (BIO2mean) et la température moyenne du trimestre le plus sec (BIO9mean), sont les moins corrélées aux autres variables et peuvent être conservées. Les autres étant fortement liées entre elles, nous choisissons de ne conserver dans le modèle initial que la température moyenne annuelle (BIO1mean), variable qui a été retenue dans le modèle de Thomaes *et al.* (2008).

Tableau 8. Corrélations de Spearman de valeurs  $|rs| \geq 0,5$  pour les variables de température  
 Les variables n'ayant aucune corrélation dépassant ce seuil ne sont pas présentées  
 Seuil de significativité : \*\*\* : 0,001

Variabiles	bio1mean	bio2mean	bio4mean	bio5mean	bio6mean	bio7mean	bio8mean	bio10mean
bio3mean		0,7***						
bio4mean	0,73***							
bio5mean	0,98***	0,52***	0,73***					
bio6mean	0,95***		0,57***	0,9***				
bio7mean	0,67***	0,85***	0,71***	0,77***				
bio8mean	0,99***		0,78***	0,98***	0,93***	0,7***		
bio10mean	0,99***		0,78***	0,98***	0,93***	0,7***	1***	
bio11mean	0,98***		0,61***	0,96***	0,97***	0,6***	0,96***	0,96***

Pour les précipitations (variables BIO12mean à BIO19mean, Tableau 9), les fortes corrélations obtenues permettent de ne retenir que le cumul annuel de précipitation (BIO12mean), également mis en avant par Thomaes *et al.* (2008).

Les nombreuses inter-corrélations qui existent pour les variables climatiques, entre elles et vis-à-vis de la topographie (altitudes uniquement) ne sont pas présentées en détail ici. On retiendra simplement que tous les descripteurs de températures sont affectés par les fortes altitudes alors qu'inversement, les précipitations augmentent et leur saisonnalité décroît.

Tableau 9. Corrélations de Spearman de valeurs  $|rs| \geq 0,5$  pour les variables de précipitation  
 Les variables n'ayant aucune corrélation dépassant ce seuil ne sont pas présentées  
 Seuil de significativité : \*\*\* : 0,001

Variables	bio12mean	bio13mean	bio14mean	bio15mean	bio16mean	bio17mean	bio18mean
bio13mean	0,93***						
bio14mean	0,97***	0,85***					
bio15mean	-0,71***	-0,49***	-0,84***				
bio16mean	0,96***	0,94***	0,88***	-0,54***			
bio17mean	0,99***	0,89***	0,99***	-0,78***	0,92***		
bio18mean	0,96***	0,94***	0,88***	-0,54***	0,99***	0,92***	
bio19mean	0,91***	0,77***	0,95***	-0,89***	0,78***	0,93***	0,78***

### 3. Résultats

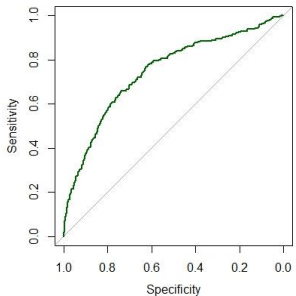
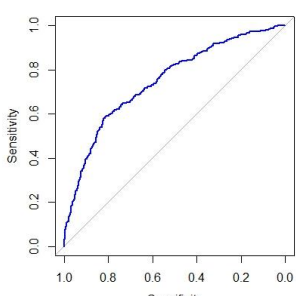
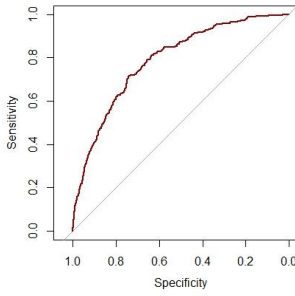
#### a) Performances des modèles

Le premier résultat issu de la modélisation est la faible puissance explicative des différents modèles. Il s'avère que les observations de Lucanes traitées ici dépendent plus fortement du hasard que de tout autre paramètre, en dehors d'un biais observateur très fort, nous y reviendrons. Les modèles ne sont pas tous évalués selon les mêmes méthodes statistiques mais leur faiblesse peut être illustrée par la performance limitée du modèle RandomForest qui n'explique que 7,02 % de la variance totale.

Cette précaution prise, nous pouvons classer les modèles entre-eux du point de vue de leurs performances au moyen d'un indice : l'AUC. L'Area Under the Curve compare les valeurs observées avec celles prédites par le modèle et prend des valeurs entre 0 et 1. Il rend ainsi compte des vrais positifs (Sensitivity dans les graphiques du tableau 10) par rapport aux faux positifs (Specificity).

On attend des valeurs d'AUC les plus proches possible de 1 et en tout cas, pour des prédictions réalistes, supérieures à 0,5 (Fawcett 2006). Les AUC obtenus ici sont corrects avec des valeurs de l'ordre de 0,75. Le plus efficace de nos modèles est le ZeroInflated construit avec toutes les variables sélectionnées (Tableau 10).

Tableau 10. Indices AUC et courbes ROC des trois modèles

Indices	RF	GLM	ZI
AUC (Area under the curve)	0,7443	0,7500	0,7878
ROC (Receiver Operating Characteristic)			

#### b) Facteurs explicatifs de la distribution du Lucane cerf-volant

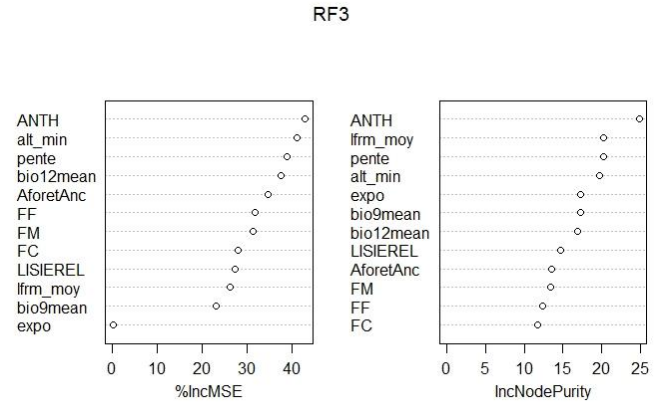
##### (1) Variables principales selon le modèle RF

La première version du modèle RF écarte clairement deux variables d'occupation du sol : d'une part les zones de fruticées (TRANS) et d'autre part les surfaces en eau (EAU). Un second modèle est donc construit sans ces deux variables, ce qui permet de comparer un total de 12 variables (Tableau 11).

La variable explicative principale est la présence de constructions (ANTH), en lien direct avec la pression d'inventaire et donc un fort biais observateur. Viennent en second plan, trois variables de topographie : le type de relief (Ifrm\_moy), la pente moyenne dans la maille et l'altitude minimale dans la maille (alt\_min). Suivent l'exposition des versants et deux variables climatiques : la température moyenne du trimestre le plus sec (bio9mean) et la précipitation totale annuelle (bio12mean). En dernier lieu se trouvent le linéaire de lisières (LISIEREL), la présence d'anciennes forêts (AforetAnc) et les trois types forestiers.

Tableau 11. Variables les plus importantes selon la modélisation RF

Variables	%IncMSE	IncNodePurity
ANTH	1,76E-02	24,85994
lfrm_moy	5,09E-03	20,25085
pente	1,18E-02	20,20913
alt_min	1,73E-02	19,75142
expo	3,08E-05	17,32722
bio9mean	6,37E-03	17,30659
bio12mean	2,08E-02	16,92859
LISIEREL	7,08E-03	14,63777
AforetAnc	1,29E-02	13,50611
FM	7,67E-03	13,42243
FF	6,81E-03	12,37781
FC	6,70E-03	11,77597



**(2) Variables principales selon le modèle GLM**

Le premier modèle retient six variables significatives (Tableau 12). Ces variables se retrouvent toutes dans la sélection effectuées par RF, en tenant compte du fait que l'indice de fermeture du paysage (IFP\_2), sélectionné par le modèle GLM, regroupe les trois types forestiers (FC, FF, FC), qui ont, eux, été retenus par le modèle RF.

Tableau 12. Variables les plus importantes selon la modélisation GLM (global)

Variables	Estimate	Std. Error	z value	Pr(> z )	Significativité
(Intercept)	2,416	1,735	1,393	0,1636	
ANTH	3,112	0,3446	9,031	<2E-016	***
IFP_2	1,027	0,2542	4,038	5,39E-05	***
bio9mean	0,03859	0,009767	3,951	7,78E-05	***
bio12mean	-0,01225	0,002247	-5,451	5,00E-08	***
expo	0,004367	0,002249	1,942	0,0522	.
LISIEREL	0,0001521	6,361E-05	2,391	0,0168	*

Au vu de la proximité entre les résultats obtenus par RF et ceux de la première version du GLM, nous avons construit un second modèle GLM avec les variables sélectionnées par RF, en incluant ainsi de facto les variables sélectionnées par le premier GLM. Il apparaît que la capacité explicative de ce second modèle GLM, plus complet, est aussi légèrement plus importante que celle du GLM initial ( $AUC_{global} = 0,7500$  contre  $AUC_{RF} = 0,7475$ ).

Les variables mises en avant par ce second GLM construit avec les variables retenues par RF, sont la présence de constructions (ANTH) et de forêts, peu importe le type. La température moyenne du trimestre le plus sec (bio9mean) vient ensuite mais participe très peu au modèle, ainsi que les variables suivantes (Tableau 13). Ces autres variables retenues, significatives mais avec des capacités explicatives très faibles, sont la précipitation annuelle (bio12mean) qui joue un rôle négatif (plus la quantité de pluie annuelle est forte, plus les chances

d’observer un Lucane sont faibles), l’exposition et l’altitude qui doit être la plus faible possible pour augmenter les chances d’observations et enfin le linéaire de lisières (LISIEREL).

Tableau 13. Variables les plus importantes selon la modélisation GLM (variables retenues par RF)

Variabiles	Estimate	Std. Error	z value	Pr(> z )	Significativité
(Intercept)	0,8227	2,014	0,408	0,682959	
ANTH	3,585	0,3927	9,128	<2E-16	***
FM	1,114	0,3313	3,363	0,00077	***
FF	1,078	0,3725	2,894	0,003804	**
FC	1,05	0,3203	3,277	0,001049	**
bio9mean	0,0363	0,01007	3,604	0,000313	***
bio12mean	-0,008749	0,003028	-2,889	0,003861	**
expo	0,004112	0,002279	1,804	0,071159	.
alt_min	-0,003341	0,001989	-1,679	0,093064	.
LISIEREL	0,0002446	6,183E-05	3,956	7,63E-05	***

### (3) Variables principales selon le modèle ZI

La première variable à être exclue de la modélisation par ZeroInflated, est l’indice de cultures et prairies (ICP2) dont les valeurs sont presque toujours nulles sauf un très petit nombre d’éléments aux valeurs très élevées. Si cela ne pose pas de problèmes dans le cas des deux précédents modèles (poids négligeable dans l’analyse), lorsqu’elle est incluse dans un modèle ZI, cette variable conduit à des résultats aberrants, en lien avec la très grande quantité de zéro qu’elle renferme.

Comme précédemment, nous avons construit deux modèles ZI avec toutes les variables d’une part et avec les variables sélectionnées par RF d’autre part. Comme avec le modèle GLM, le modèle le plus complet est aussi le plus solide, et a donc été retenu ( $AUC_{global} = 0,7878$  contre  $AUC_{RF} = 0,7710$ ).

La modélisation ZI se fait en deux étapes de calcul. La première conduit à peu de chose près aux mêmes résultats que les analyse précédentes, où l’on retrouve la présence incontournable de ANTH au coté d’autres variables dont les types forestiers et d’autres variables à poids faible dans cette partie de l’analyse. La seconde étape, prend en compte des zéros en surnombre (absences d’observations liées à des facteurs autres que l’absence réelle du Lucane), ce qui change complètement la donne, avec en premier lieu l’élimination de la variable ANTH comme cela était envisagé (Tableau 14).

Les variables finalement retenues sont celles dont les niveaux de significativité sont améliorés par la prise en compte des zéros en surnombre. Il s’agit de la présence de vergers (RG), de chênaie âgée (Achene), d’espaces protégés (Aprotect) et de l’altitude (alt\_min). Étant donné le mode de calcul utilisé ici, focalisé sur les absences d’observations, les signes des contributions des variables (Estimate) sont inversés par rapport à ce que nous avons vu pour le modèle GLM qui prend en compte les observations. Selon le modèle ZI, les observations de Lucane sont donc plus fréquentes lorsqu’il y a moins de vergers, plus de chêne et d’espaces protégés ainsi que des altitudes plus basses.



Tableau 14. Variables les plus importantes selon la modélisation ZI (global)

Variables	Count					Zero-inflation				
	Estimate	Std. Error	z value	Pr(> z )	Significativité	Estimate	Std. Error	z value	Pr(> z )	Significativité
Log(theta)	11,54	65,95	0,175	0,86108						
(Intercept)	-0,06633	4,857	-0,014	0,9891		-40,16	36,52	-1,099	0,2716	
ANTH	2,785	0,3889	7,159	8,10E-13	***	-21,75	17,89	-1,216	0,224	
FM	1,462	0,6628	2,206	0,02736	*	18,21	11,19	1,627	0,1036	
RG	1,289	0,6957	1,853	0,06387	.	18,06	8,074	2,236	0,0253	*
FC	1,125	0,6397	1,759	0,07862	.	15,89	11,19	1,419	0,1558	
FF	0,7064	0,6941	1,018	0,30885		13,91	10,94	1,272	0,2035	
IFP_2	0,6121	0,3656	1,674	0,09412	.	-4,601	8,222	-0,56	0,5758	
lfrm_moy	0,05889	0,1206	0,488	0,62541		0,424	0,4579	0,926	0,3544	
bio2mean	-0,03882	0,05005	-0,776	0,43797		-0,2338	0,2735	-0,855	0,3927	
bio1mean	0,03273	0,03533	0,926	0,35424		0,3659	0,2376	1,54	0,1236	
Achene	-0,02687	0,01028	-2,615	0,00893	**	-0,2052	0,08035	-2,554	0,0106	*
bio9mean	0,02396	0,01049	2,283	0,02244	*	-0,000572	0,07186	-0,008	0,9937	
bio12mean	-0,008193	NA	NA	NA		0,003481	0,02104	0,165	0,8686	
penne	0,008006	0,01111	0,721	0,4712		0,05262	0,05185	1,015	0,3102	
expo	0,00383	0,002255	1,698	0,08946	.	0,01296	0,01223	1,06	0,2892	
AforetAnc	0,001989	0,003769	0,528	0,59767		-0,003066	0,01799	-0,17	0,8647	
Aprotect	-0,001486	0,001504	-0,988	0,32298		-0,1084	0,06181	-1,754	0,0794	.
alt_min	0,001307	0,003081	0,424	0,67143		0,04197	0,01821	2,304	0,0212	*
LISIEREL	2,122E-05	5,986E-05	0,354	0,723		-0,000411	0,000546	-0,752	0,452	

### c) Prédiction de présence du *Lucane cerf-volant*

Nous venons de dégager les facteurs environnementaux qui expliquent la distribution du Lucane. En retour, nous pouvons utiliser les possibilités de prédiction des modèles pour établir des distributions théoriques de Lucane. Sachant les variables impliquées et leurs contributions dans chaque modèle, et disposant des valeurs de chacune de ces variables sur toute la carte analysée, il est en effet possible de demander au modèle de rechercher les mailles où les facteurs environnementaux sont favorables à la présence du Lucane. Trois cartes de prédictions ont ainsi été créées (Illustrations 7, 8, 9).

Les prédictions des trois modèles sont relativement similaires. Le modèle RF propose beaucoup plus de mailles où le Lucane pourrait être présent. Les modèles GLM et ZI, plus restrictifs, jouent un rôle de filtre pour ne finalement retenir que 423 mailles prédisant la présence du Lucane cerf-volant (Illustration 10).

Carte p1. Prédiction de présence de Lucane cerf-volant : RandomForest

Légende

Predict Lucane RF

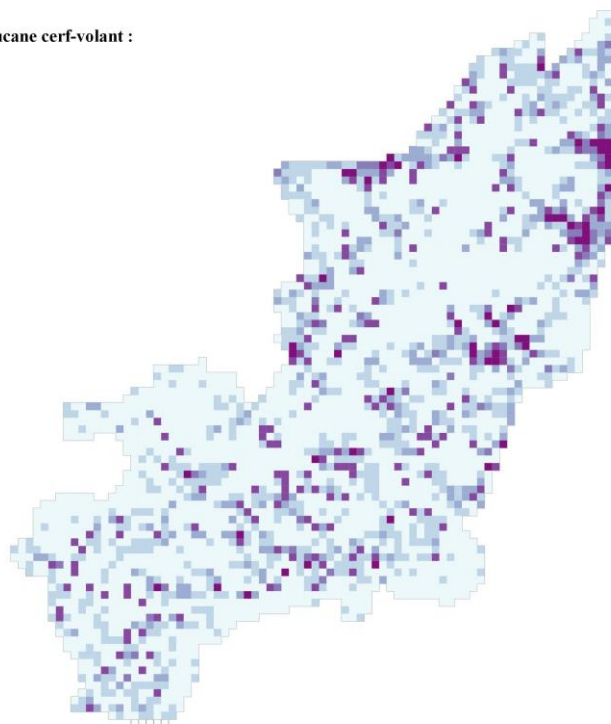
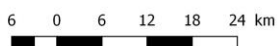
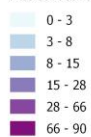


Illustration 7. Prédiction de présence du Lucane cerf-volant selon le modèle RF

Carte p2. Prédiction de présence de Lucane cerf-volant : GLM

Légende

Predict Lucane GLM

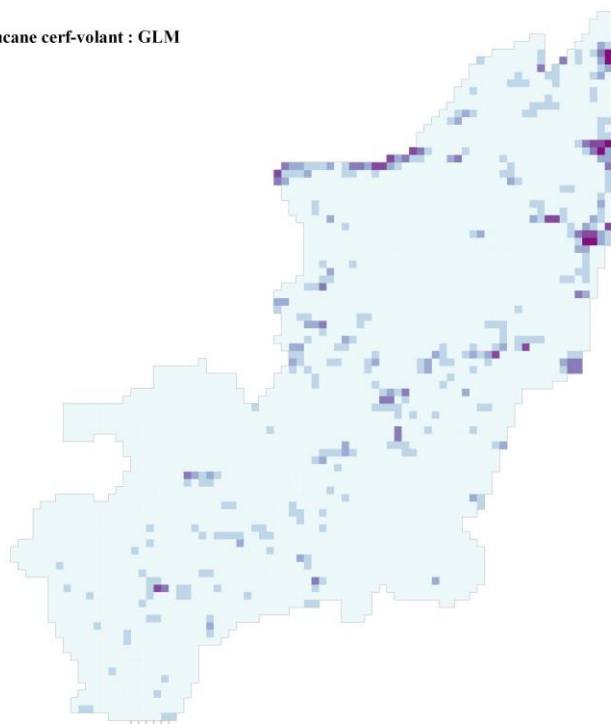
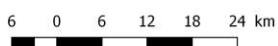
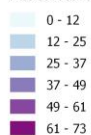


Illustration 8. Prédiction de présence du Lucane cerf-volant selon le modèle GLM

---

*Illustration 9. Prédiction de présence du Lucane cerf-volant selon le modèle ZI*

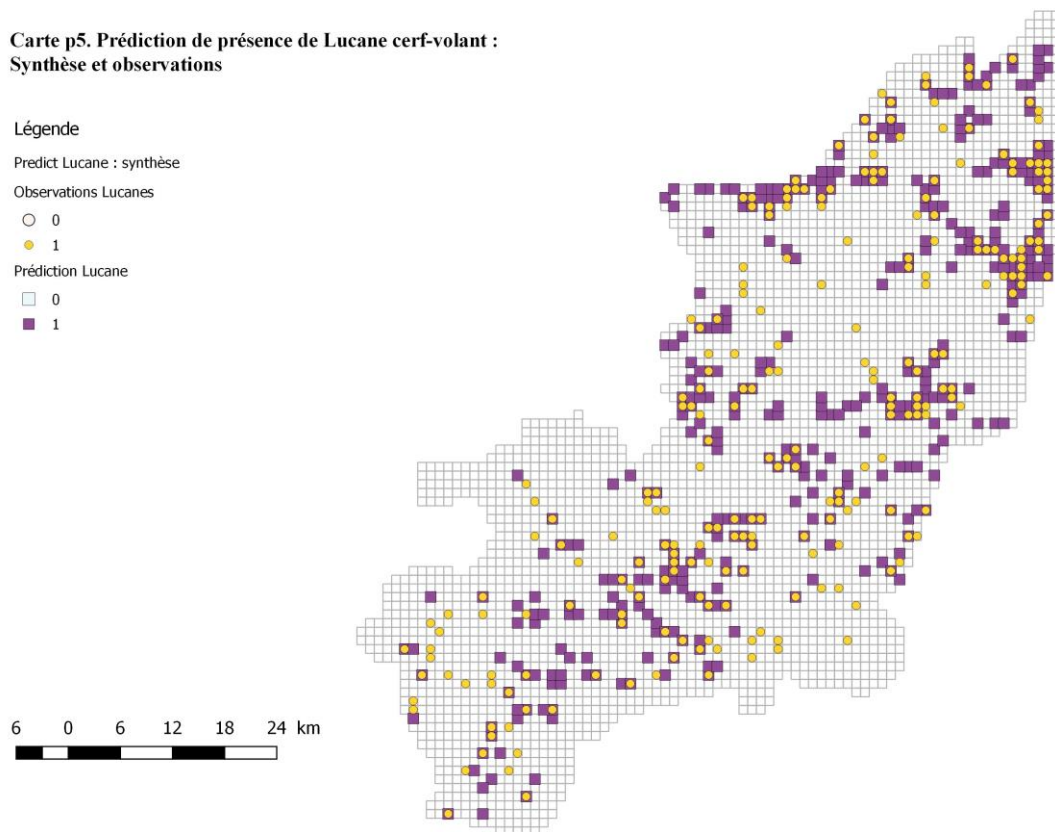
---

---

*Illustration 10. Prédiction de présence du Lucane cerf-volant : synthèse des 3 prédictions*

---

La plupart des prédictions recourent les distributions mesurées, en particulier avec le modèle RF. Les prédictions issues de la synthèse des trois modèles sont donc, comme attendu, très significativement corrélées aux observations (Spearman  $r_s = 0,41402$ ,  $p < 0,0001$ ) bien que 93 mailles avec observations, sur 242, ne soient pas prédites (38,4%, Illustration 11).



*Illustration 11. Prédiction de présence du Lucane cerf-volant : synthèse des 3 prédictions et observations réelles*

#### ***d) Analyse d'après les prédictions de présence du Lucane***

##### ***(1) Prédications et variables "de gestion"***

Cinq des variables recouvrent des éléments sur lesquels il est possible d'agir ou de construire des politiques de gestions : ANTH, AforetAnc, LISIEREL, Aprotect et Achene. Il est donc important d'examiner les relations entre ces variables liées à la gestion et les résultats de la prédiction de distribution du Lucane cerf-volant (Tableau 15).

ANTH, LISIEREL, AforetAnc et Achene ont été mises en évidence par les modèles, particulièrement le modèle ZI pour les deux dernières. Il serait cohérent qu'elles prennent des valeurs différentes selon qu'elles se trouvent dans un espace où le Lucane est prédit ou dans un espace où le Lucane n'est pas prédit. C'est ainsi le cas avec ANTH et LISIEREL qui affichent des valeurs supérieures lorsque le Lucane est prédit. Cela correspond bien aux résultats précédents.

En revanche, AforetAnc et Achene prennent des valeurs significativement moins fortes lorsque le Lucane est

prédit, ce qui va à l’encontre des résultats du modèle ZI. La forte emprise de la variable ANTH sur les résultats des modèles RF et GLM, et donc au final sur les prédictions, vient expliquer cette situation, ANTH étant inversement corrélée à AforetAnc (Spearman  $r_s = -0,48694$ ,  $p < 0,0001$ ) et plus faiblement à Achene (Spearman  $r_s = -0,22713$ ,  $p < 0,0001$ ).

Les valeurs de Aprotect ne sont pas différentes selon le statut de prédiction de présence de Lucane. L’ajout de surfaces supplémentaires, sans création d’entités nouvelles (scénario 1) ne modifie pas ce résultat, et ce même avec 20 % supplémentaires. Cette variable ressortait comme quasiment significative dans le modèle ZI. Dans leur configuration spatiale actuelle, et avec les données d’observations dont nous disposons, les espaces protégés paraissent ainsi ne pas jouer de rôle vis-à-vis du Lucane.

Nos prochaines étapes d’analyse vont questionner la protection du Lucane en fonction d’accroissements des surfaces protégées, autour de l’existant puis en plaçant de nouveaux espaces selon divers scénarios.

Tableau 15. Relations entre les variables liées à la gestion selon la prédiction de présence de Lucane  
*H<sub>0</sub>* : les valeurs des variables sont équivalentes que le Lucane soit prédit ou non

Variable	Valeur moyenne		F	PERMANOVA	
	Prédiction Lucane = 0	Prédiction Lucane = 1		p	Conclusion
ANTH	0,038	0,265	1150	0,0001	Rejet H <sub>0</sub>
AforetAnc	59,237	47,289	26,71	0,001	Rejet H <sub>0</sub>
Aprotect	11,646	9,846	0,3784	0,546	Non rejet H <sub>0</sub>
Aprotect_5	14,475	12,098	0,3784	0,523	Non rejet H <sub>0</sub>
Aprotect_10	15,164	12,674	0,3784	0,532	Non rejet H <sub>0</sub>
Aprotect_20	16,543	13,826	0,3784	0,577	Non rejet H <sub>0</sub>
Achene	5,091	2,150	31,34	0,001	Rejet H <sub>0</sub>
Achene_5	6,242	3,075	31,34	0,001	Rejet H <sub>0</sub>
Achene_10	6,540	3,221	31,34	0,001	Rejet H <sub>0</sub>
Achene_20	7,134	3,514	31,34	0,001	Rejet H <sub>0</sub>
LISIEREL	760,817	1717,387	319,4	0,001	Rejet H <sub>0</sub>
LISIEREL_5	798,858	1803,257	319,4	0,001	Rejet H <sub>0</sub>
LISIEREL_10	836,899	1889,126	319,4	0,001	Rejet H <sub>0</sub>
LISIEREL_20	912,980	2060,865	319,4	0,001	Rejet H <sub>0</sub>

## (2) Prédiction et scénarios

Cette analyse exploratoire cherche à mesurer des différences entre les prédictions de présence de Lucane calculées par le meilleur modèle : d’abord à partir des données initiales puis des données modifiées artificiellement selon les différents scénarios.

De nouveau, le biais d’échantillonnage très important en faveur de la variable ANTH, ne permet pas d’obtenir de résultats probant en modifiant des variables sur lesquelles des politiques de gestion pourraient s’appuyer. Nous présentons néanmoins une illustration du procédé, à titre d’exemple, avec les surfaces d’aires protégées, la présence de chênaies matures et la longueur de lisières.

Pour toutes les variantes artificielles, le modèle ZI ne construit pas de prédictions significativement différentes lorsque les changements liés aux scénarios sont appliqués (Tableau 16 et test de Kruskal-Wallis  $\chi^2 = 3516$ ,  $p = 0,4968$  pour toutes les configurations du tableau). Ce résultat est illustré pour l’un des scénario impliquant

les espaces protégés, par l', où aucune différence d'ampleur n'est décelable entre les prédictions issues des mesures et les prédictions issues de la variable artificielle. Nous n'avons retenu qu'un exemple mais il en va de même avec toutes les autres configurations testées (non présentées).

Dans le modèle ZI, les variables testées (Aprotect, Achene, LISIEREL) ont des poids négligeables face à ANTH ce qui ne permet pas de tester quoi que ce soit au niveau de ces variables, sauf à augmenter de façon sur-réaliste leurs valeurs (20 % étant déjà une valeur très importante !).

Pour compléter notre exemple, nous avons appliqué la même analyse au modèle GLM retenu (cf. p. 23, modèle auquel nous avons ajouté Aprotect), avec la variante de scénario : 10 entités réparties aléatoirement pour +10 % de surface protégée. À l'inverse des résultats obtenus avec le modèle ZI, les deux prédictions obtenues sont significativement différentes (Student :  $t = -310,73$ ,  $p < 2,2e-16$  ; Kruskal-Wallis  $\chi^2 = 3516$ ,  $p = 0,03611$  ; voir l'). Nous retrouvons ici le biais d'échantillonnage qui ne permet pas de statuer de façon fiable. La présence de nombreuses absences d'observations dans les données initiales (faux zéro), opposent complètement les résultats des modèles selon qu'ils en tiennent compte (ZI) ou non (GLM).

Tableau 16. Relations entre les prédictions de présence de *Lucane* calculées par le modèle ZI à partir des données initiales puis des données de scénario

$H_0$  : la "prédiction initiale" est équivalente à la "prédiction scénario"

Scénarios	Test de Student		Conclusion
	t	p	
Aprotect : toutes variantes de scénario	0	1	Non rejet $H_0$
Achene_5	0,067778	0,946	Non rejet $H_0$
Achene_10	0,13738	0,8907	Non rejet $H_0$
Achene_20	0,28084	0,7788	Non rejet $H_0$
LISIEREL_5	-0,077197	0,9385	Non rejet $H_0$
LISIEREL_10	-0,15415	0,8775	Non rejet $H_0$
LISIEREL_20	-0,30735	0,7586	Non rejet $H_0$

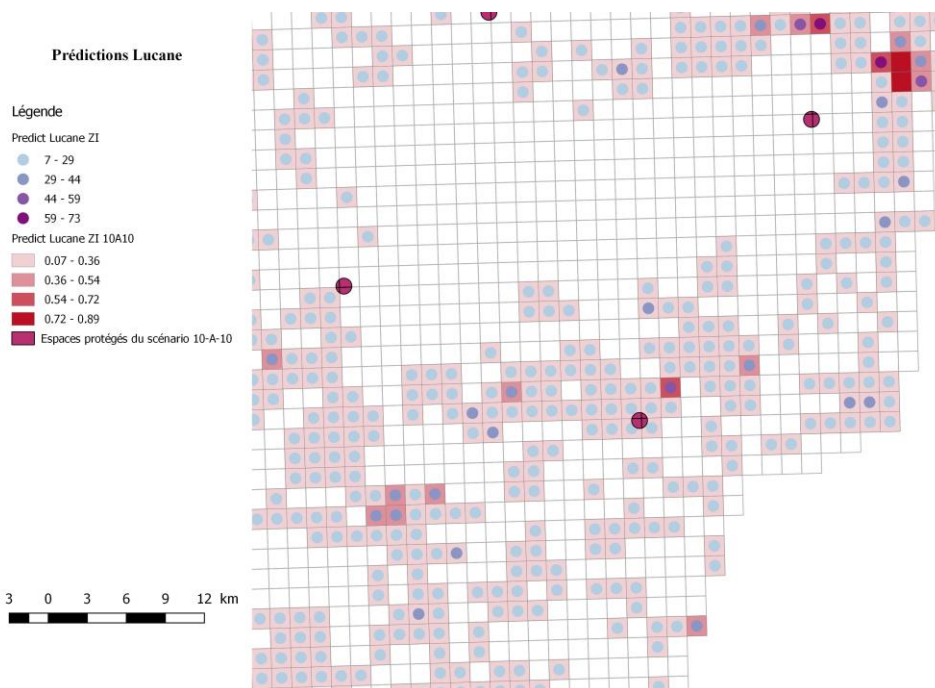


Illustration 12. Prédictions de présence de Lucane par le modèle ZI basé sur les données de terrain et sur les données modifiées du scénario : 10 grandes entités réparties aléatoirement pour +10 % de surface protégée (détail)

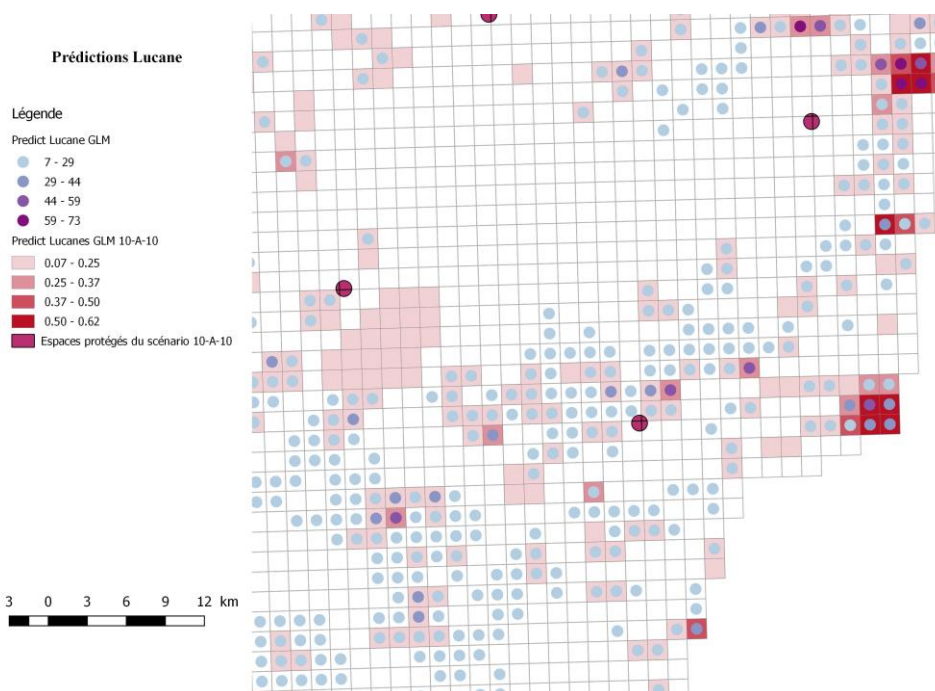


Illustration 13. Prédictions de présence de Lucane par le modèle GLM basé sur les données de terrain et sur les données modifiées du scénario : 10 grandes entités réparties aléatoirement pour +10 % de surface protégée (détail)

### **(3) Prédiction et réseaux d'îlots de sénescence**

À l'issue de leur étude du réseau d'îlots de sénescence coté français, Demets et Bruciamacchie (2017) ont proposé deux options pour la créations d'un réseau d'îlots, selon deux méthodes de calcul : Simple (D1) et Electre (D2) (). Ces réseaux représentent 450,64 ha pour le réseau D1 et 458,65 ha pour le réseau D2. Le réseau actuel couvre 216,17 ha et apparaît comme non relié statistiquement aux prédictions de présence du Lucane (Tableau 17).

Nous avons testé les interactions entre les réseaux d'îlots proposés (D1 et D2) et nos prédictions de présence de Lucane. Deux variables ont été explorées : la surface d'îlots par km<sup>2</sup> (par maille) puis la surface d'îlot dans un espace de 8 mailles réparties autour de chaque maille, soit un rayon de 1,5 km autour du centre de chaque maille (buffer de 8 km<sup>2</sup>).

*Illustration 14. Réseaux d'îlots de sénescence du PNRVN : actuel et proposés par Demets et Bruciamacchie (2017)*

Le réseau d'îlots D1 est indépendant de la prédiction de présence du Lucane. Les valeurs non significatives du test ne permettent pas de différencier les surfaces proposées pour le réseau D1 selon que le Lucane soit prédit ou non.

En revanche, le réseau D2, et à sa suite la fusion des réseaux D1 et D2, sont significativement différents selon la prédiction de présence du Lucane. Les surfaces proposées pour le réseau d'îlot D2 sont plus importantes lorsque le Lucane est prédit. Elles sont ainsi en cohérence avec une politique de protection des habitats du Lucane.



Tableau 17. Relations entre chaque variable descriptive des réseaux d'îlots proposés par Demets et Bruciamacchie pour le PNR des Vosges du Nord, selon la prédiction de présence de Lucane  
 $H_0$ : les valeurs des variables sont équivalentes que le Lucane soit prédit ou non

Variable	Surface moyenne (ha)		F	PERMANOVA	
	Prédiction Lucane = 0	Prédiction Lucane = 1		p	Conclusion
Îlots existants : SEN 2016 PNRVN	0,1406	0,1366	0,005193	0,949	Non rejet $H_0$
Îlots Simple D1 : Aire (km <sup>2</sup> )	0,2861	0,3726	0,4799	0,496	Non rejet $H_0$
Îlots Simple D1 : Aire (Buffer 8 km <sup>2</sup> )	0,9009	1,0316	0,4507	0,507	Non rejet $H_0$
Îlots Electre D2 : Aire (km <sup>2</sup> )	0,2763	0,5670	4,309	0,033	Rejet $H_0$
Îlots Electre D2 : Aire (Buffer 8 km <sup>2</sup> )	0,9863	1,8580	12,07	0,001	Rejet $H_0$
Îlots D1+D2 : Aire (km <sup>2</sup> )	0,5117	0,8604	3,762	0,065	Non rejet $H_0$
Îlots D1+ D2 : Aire (Buffer 8 km <sup>2</sup> )	1,5094	2,9535	23,23	0,001	Rejet $H_0$

## 4. Conclusions

Cette étude couvre le territoire de la Réserve de Biosphère Transfrontalière des Vosges du Nord - Pfälzerwald. L'objectif général est de déterminer si cet espace garantit les conditions nécessaires à l'existence des espèces saproxyliques et en particulier le lucane cerf-volant, espèce emblématique.

Nous avons rassemblé un important jeu de données regroupant des variables environnementales et des observations de Lucane. Le volet traitant des coléoptères indicateurs, autres que le Lucane, a dû être écarté faute de disposer, dans des délais et charges de travail raisonnables, des données adaptées à nos besoins.

Pour le Lucane, nous avons mis en place des méthodes d'analyses à même de décrire ses habitats favorables (variables environnementales et anthropiques, incluant l'aménagement forestier), puis dans un second temps d'évaluer différents scénarios de gestion. Comme cela a été souligné et sera discuté plus détail ci-dessous, de meilleures observations permettraient de fournir des modèles plus puissants. En conséquence de ce travail à partir de données biaisées, nos conclusions ne sont pas suffisamment robustes pour orienter avec certitude vers des démarches de gestion. Néanmoins, la qualité et la quantité de données ont été assez solides pour élaborer une démarche d'analyse visant à remplir ces objectifs.

Les modèles et les scénarios que nous proposons ont montré qu'il est possible d'une part, d'évaluer l'environnement favorable aux Lucanes et d'autre part, d'évaluer des mesures d'aménagement favorisant la présence de cette espèce. Nous avons montré que d'importantes applications sont possibles bien que nos tests doivent en l'état, être considérés comme des exemples de démarche d'évaluation. Ces applications permettent de tester en détail différents scénarios d'aménagement en faveur du Lucane, par exemple en augmentant la surface de protection selon une répartition aléatoire ou attenante aux zones déjà existantes. Notre étude s'est tournée vers les surfaces d'espaces protégés et les îlots de sénescence proposés sur le territoire français. Elle pourrait tout autant se baser sur la proportion d'une essence dans le paysage, l'ouverture du boisement ou encore les longueurs de lisières. Le principal écueil qu'il nous faut à présent résoudre n'est pas statistique mais lié aux données d'observation de Lucane.

Pour illustrer le biais présent dans les données d'observation, il faut rappeler que l'écart est d'un facteur 100 entre la participation de la variable anthropique (ANTH) et la variable significative suivante, la surface de chênaie mature (Achene) dans le modèle que nous avons sélectionné (méthode ZeroInflated). Il s'agit d'un biais observationnel habituel, très régulièrement présent dans toutes les études de science participative. Par exemple, ce même biais a été mis en évidence dans l'étude du Lucane conduite dans le sud de l'Angleterre par Percy *et al.* (1999). Ces auteurs concluent que la présence des Lucanes est reliée aux jardins dans lesquels les observateurs passent beaucoup de temps. Pour Tiago *et al.* (2017), la densité des chemins et la présence de lacs (auprès desquels le public séjourne) sont les variables les plus pertinentes pour expliquer les différentes observations réalisées à l'échelle du Portugal pour différentes espèces.

Il faut de fait considérer la variable ANTH comme un reflet de la densité des observateurs. Ainsi, la carte des données de Lucane traduit les contacts entre le Lucane et les observateurs et ne reflète pas la présence réelle de l'espèce. La nuance est fondamentale vis-à-vis de nos questionnements en matière de gestion et de la manière d'y répondre.

Les données d'observations de Lucanes se caractérisent donc par la présence d'un très grand nombre de zéros. Ces zéros en surnombre répondent de quatre cas (Zuur *et al.* 2009) :

- (1) Une source liée au contexte : l'habitat n'est pas favorable, le Lucane est absent. Cette première source de zéro intègre de fait les absences d'observations en dehors des habitats du Lucane ;
- (2) Ils découlent de l'échantillonnage qui n'a pas été effectué dans les conditions nécessaires pour détecter le Lucane ;
- (3) Ils sont liés aux compétences des observateurs dont certains ont pu confondre le Lucane cerf-volant avec par exemple le *Dorcus* ou de gros *Carabus* noirs comme nous l'avons constaté ;
- (4) L'absence est due au Lucane lui-même qui n'est pas présent alors que l'habitat est favorable.

Les données relevant du 1<sup>er</sup> cas peuvent être retirées en amont des analyses car elles ne correspondent pas à la problématique de l'espèce. Ce sont par exemple les espaces en eau où le Lucane ne peut pas vivre. Le 4<sup>e</sup> cas rassemble les vrais négatifs qui sont les situations importantes à comprendre et à analyser : pourquoi le Lucane est absent alors que la situation lui est favorable ? A-t-il disparu dans le passé ? Les situations 2 et 3 sont des faux négatifs qui doivent être corrigés du mieux possible. Nos données relèvent surtout d'une variante du cas 2 avec des prospections qui n'ont pas concerné suffisamment certains habitats. Elles n'ont pas été homogènes spatialement. Jusqu'à quel point ces faux-négatifs sont présents dans les données d'observations ?

Ce taux de faux-négatifs n'a pas pu être estimé dans notre étude car il faut pour cela connaître le plan d'échantillonnage, c'est à dire les conditions dans lesquelles sont obtenues les observations et surtout les absences d'observations, au moins dans une partie des sites étudiés (Tyre *et al.* 2003, Guillera-Arroita 2016). Il est donc nécessaire de rassembler des données relatives au mode d'échantillonnage. A minima, il s'agit de connaître les sites qui ont été prospectés mais où aucune observation n'a été réalisée malgré des conditions favorables (mode de recherche, période de l'année). Il est également important de pouvoir disposer de métadonnées décrivant les observateurs : quelle est leur connaissance du Lucane et leur capacité à le détecter ? Quel est le nombre de sorties effectuées avant de détecter un premier Lucane ? etc. (pour plus de détails voir par exemple Bird *et al.* 2014 ou Guillera-Arroita 2016). Ces métadonnées permettraient d'affiner les modélisations en ouvrant la possibilité d'estimer un taux de détection pour chaque observateur ainsi qu'un taux de détectabilité pour le Lucane.

Dans le même ordre d'idée, un travail cartographique pourrait être effectué pour déterminer la facilité de prospection des espaces. Certains, par exemple moins accessibles, verraient leurs données d'observations renforcées dans le modèle par rapport à des sites plus accessibles où les Lucanes seraient vus plus régulièrement simplement du fait de la plus forte fréquentation. La facilité de prospection pourrait être reflétée par la densité de chemins, les données de fréquentation des massifs forestiers ou encore la distance aux parkings les plus proches. Ces données cartographiques nous font défaut actuellement mais peuvent être facilement construites.

L'effort principal à fournir à présent est celui d'un échantillonnage plus complet, aléatoire idéalement, visant une prospection sur tous les types d'occupation du sol et de statuts (protégé, forêt ancienne, etc.). Les espaces protégés paraissent en effet actuellement ne jouer qu'un rôle mineur vis-à-vis du Lucane, peut-être du fait d'un défaut de prospection. Le taux global d'observation dont nous disposons est de 6,89 % pour les deux parcs. Pour comparaison, il est d'environ 20 % pour la Belgique dans l'étude de Thomaes *et al.* (2008) et de 63 % dans le suivi des papillons de Belgique (Maes *et al.* 2003).

Pour des raisons pratiques de faisabilité (nombre d'observateurs, accessibilité du terrain...), cette nouvelle phase d'échantillonnage ne pourra certainement pas couvrir un grand nombre de mailles. Il semble raisonnable de la limiter en deçà des 13 % qui nous séparent de l'étude belge (soit 457 mailles à prospecter). Il est ainsi possible de calibrer les modèles sur de petites séries de mailles bien échantillonnées puis d'appliquer les modèles au jeu de données complet. Cette stratégie est classiquement recommandée pour des problématiques comme la nôtre (Maes *et al.* 2003, MacKenzie 2005).

De plus, le fait de répéter les mesures sur une petite série de maille permettra de mieux calculer le taux de détectabilité du Lucane. Il s'agit d'une information qui n'est pas accessible avec de simples observations ponctuelles alors qu'elle permet de consolider la construction des modèles (Guillera-Arroita 2016). Pour alléger le temps de prospection, la répétition des observations peut être stoppée dès que le Lucane est détecté une première fois dans la maille.

Des réseaux d'observateurs de Lucane ayant été mis en place au cours de notre étude, ils pourraient contribuer aux nouvelles campagnes de prospection, cadrées cette fois par un protocole comme cela vient d'être présenté. Ainsi, il serait possible de réaliser des sessions d'observations impliquant plusieurs personnes au sein de zones géographiques limitées mais qui seront bien prospectées (dates optimales, même contexte climatique, temps de prospection connu, etc.). Ces secteurs devront être choisis loin des villes et villages (>2 km) et devront inclure des espaces protégés et des contextes paysagers variés. Cela permettra en plus, de créer une animation de type évènementiel autour de ces recherches et de communiquer sur le programme Lucane. Il serait intéressant également de tester l'emploi de chiens renifleurs dressés à détecter les larves de Lucane. De tels chiens sont utilisés avec succès pour l'osmoderne (voir par exemple Mosconi *et al.* 2017), et les nids de tortue marines. L'équipe de A. Thomaes a semble-t-il conduit quelques essais pour le Lucane (A. Camparano com. pers. janv. 2019). L'usage de chiens permettrait de résoudre en partie la question des mauvaises détections et serait un gain de temps important sur le terrain.

## 5. Bibliographie

- Alexander KNA (2008) Tree biology and saproxylic coleoptera: issues of definitions and conservation language, *Revue d'Ecologie (Terre Vie)*, 63, 1–7.
- Bird TJ, Bates AE, Lefcheck JS, Hill NA, Thomson RJ, Edgar GJ, Stuart-Smith RD, Wotherspoon S, Krkosek M, Stuart-Smith JF, Pecl GT, Barrett N, Frusher S (2014) Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation* 173: 144–154. doi: 10.1016/j.biocon.2013.07.037
- Brustel H (2001) *Coléoptères saproxyliques et valeur biologique des forêts françaises*. Institut National Polytechnique
- Cutler A, Cutler DR, Stevens JR (2012) *Random Forests*. In: Zhang C, Ma Y (Eds), Ensemble Machine Learning. Springer US, Boston, MA, 157–175. doi: 10.1007/978-1-4419-9326-7\_5
- Demets V, Bruciamacchie M (2017) *Audit du réseau actuel d'îlots de sénescence Propositions d'évolution. État au printemps 2016*. AgroParisTech, Laboratoire d'Économie Forestière, Nancy, 60 pp.
- Dodelin B (2010) Les Coléoptères saproxyliques, derniers maillons de la forêt. *Bulletin mensuel de la Société linnéenne de Lyon* Hors-série n°2: 159–166.
- Dodelin, B., Fantino, G. et Darmon, G. 2018. *Analyse du territoire et formulation de propositions d'aménagements en faveur des espèces saproxyliques, notamment le Lucane cerf-volant, Lucanus cervus, intégrés au Programme LIFE Biocorridors - 14 NAT/FR/000290*. Rapport d'étude intermédiaire pour le PNR des Vosges du Nord et le Naturpark Pfälzerwald, 45 pp.
- Eckelt A, Müller J, Bense U, Brustel H, Bußler H, Chittaro Y, Cizek L, Frei A, Holzer E, Kadej M, Kahlen M, Köhler F, Möller G, Mühle H, Sanchez A, Schaffrath U, Schmidl J, Smolis A, Szallies A, Németh T, Wurst C, Thorn S, Christensen RHB, Seibold S (2017) “Primeval forest relict beetles” of Central Europe: a set of 168 umbrella species for the protection of primeval forest remnants. *Journal of Insect Conservation*: 1–14. doi: 10.1007/s10841-017-0028-6
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27: 861–874. doi: 10.1016/j.patrec.2005.10.010
- Glaser FF, Hauke U (2004) *Historisch alte Waldstandorte und Hudewälder in Deutschland: Ergebnisse bundesweiter Auswertungen*. Bundesamt für Naturschutz, Bonn, 193 pp + CD-ROM.
- Guillera-Arroita G (2016) Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography*. doi: 10.1111/ecog.02445
- Hammer Ø (2017) *PAST - Palaeontological Statistics, version 3.18. Reference manual*. Natural History Museum & University of Oslo, Oslo, 259 pp.
- James G, Witten D, Hastie T, Tibshirani R eds. (2013) *An introduction to statistical learning: with applications in R*. Springer, New York, 426 pp.
- MacKenzie DI (2005) Was it there? Dealing with imperfect detection for species presence/absence data. *Australian & New Zealand Journal of Statistics* 47: 65–74. doi: DOI: 10.1111/j.1467-842X.2005.00372.x
- Manel S, Dias J-M, Ormerod SJ (1999) Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling* 120: 337–347. doi: 10.1016/S0304-3800(99)00113-1
- Manel S, Williams HC, Ormerod SJ (2002) Evaluating presence-absence models in ecology: the need to account for prevalence: Presence-absence modelling. *Journal of Applied Ecology* 38: 921–931. doi: 10.1046/j.1365-2664.2001.00647.x
- Mosconi F, Campanaro A, Carpaneto GM, Chiari S, Hardersen S, Mancini E, Maurizi E, Sabatelli S, Zauli A, Mason F, Audisio P (2017) Training of a dog for the monitoring of *Osmoderma eremita*. *Nature Conservation* 20: 237–264. doi: 10.3897/natureconservation.20.12688
- Dodelin, Fantino, Darmon 2018. *Lucanus cervus dans le PNR des Vosges du Nord - Pfälzerwald*

- Müller J, Bußler H, Bense U, Brustel H, Flechtner G, Fowles A, Kahlen M, Möller G, Mühle H, Schmidl J, Zabransky P (2005) Urwald relict species - Saproxyllic beetles indicating structural qualities and habitat tradition. *Waldoekologie online* 2: 106–113.
- Percy C, Bassford G, Keeble V (1999) *Stag Beetles. Findings of the 1998 National Survey*. London: People's Trust for Endangered Species, 55 pp.
- Perez-Jaume S, Skaltsa K, Pallarès N, Carrasco JL (2017) ThresholdROC: Optimum Threshold Estimation Tools for Continuous Diagnostic Tests in R. *Journal of Statistical Software* 82. doi: 10.18637/jss.v082.i04
- Saurat, R. 2017. *Habitats potentiels du coléoptère Lucane cerf-volant (Lucanus cervus) dans la réserve de Biosphère franco-allemande transfrontalière des Vosges du Nord - Pfälzerwald*. Mémoire de Projet Tutoré de première année, Master "Sciences de l'Environnement Terrestre" spécialité Biologie, Ecologie Evolution 1ère année – TC3. spécialité Biologie, Ecologie Evolution, Aix-Marseille université, 18 pp.
- Sebek P, Barnouin T, Brin A, Brustel H, Dufrêne M, Gosselin F, Meriguet B, Micas L, Noblecourt T, Rose O, Velle L, Bouget C (2012) A test for assessment of saproxyllic beetle biodiversity using subsets of “monitoring species.” *Ecological Indicators* 20: 304–315. doi: 10.1016/j.ecolind.2012.02.033
- Thomaes A (2008) *Onderzoek en monitoring van het Vliegend hert [Study and Monitoring of the Stag Beetle]*. Instituut Natuur- en Bosonderzoek, Brussel, 24 pp. Available from: [https://data.inbo.be/purews/files/719538/Thomaes\\_2008\\_OnderzoekMonitoringVliegendHert.pdf](https://data.inbo.be/purews/files/719538/Thomaes_2008_OnderzoekMonitoringVliegendHert.pdf).
- Thomaes A (2009) *A protection strategy for the stag beetle (Lucanus cervus, (L., 1758), Lucanidae) based on habitat requirements and colonisation capacity*. In: Proceedings of the 5th Symposium and Workshop on the Conservation of Saproxyllic Beetles. Pensoft Publishers, 149–160.
- Thomaes A, Cammaerts R, Kervynz T, Beck O, Crevecoeur L (2010) Distribution and site preferences of the stag beetle, *Lucanus cervus* in Belgium (Coleoptera: Lucanidae). *Bulletin SRBE/KBVE* 146: 33–46.
- Thomaes A, Kervyn T, Maes D (2008) Applying species distribution modelling for the conservation of the threatened saproxyllic Stag Beetle (*Lucanus cervus*). *Biological conservation* 141: 1400–1410.
- Tiago P, Ceia-Hasse A, Marques TA, Capinha C, Pereira HM (2017) *Spatial distribution of citizen science casuistic observations for different taxonomic groups*. Scientific Reports 7. doi: 10.1038/s41598-017-13130-8
- Tyre AJ, Tenhumberg B, Field SA, Niejalke D, Parris K, Possingham HP (2003) Improving Precision and Reducing Bias in Biological Surveys: Estimating False-Negative Error Rates. *Ecological Applications* 13: 1790–1801. doi: 10.1890/02-5078
- Zeileis A, Kleiber C, Jackman S (2008) Regression Models for Count Data in R. *Journal of Statistical Software* 27: 1–25.
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM eds. (2009) *Mixed effects models and extensions in ecology with R*. Springer, New York, NY, 574 pp.

## 6. Annexe 1. Liste des contacts et réunions de travail

Tableau 18. Liste des contacts

Nom	Organisation	Notes	Contact
Francis Matt		Expertise coléoptères saproxyliques	francismatt@wanadoo.fr
Michel Loubère		Données de coléoptères	malachus@orange.fr
Valentin Demets	AgroParisTech, Nancy	Géo-traitement des données forêt du PNR des Vosges du Nord	demets.valentin@gmail.com
Remy Saurat	Benoît Dodelin	Expertise coléoptères saproxyliques Stage de Master 2 à partir des données Allemandes de Lucane	remy-saurat@hotmail.fr
Ulrich Diehl	Biosphärenhaus Fischbach	Spécialiste d'espèces	diehl@biosphaerenhaus.de
Helmut Schuler	Biosphärenreservat Pfälzerwald	Coordinateur du projet Hirtenwege	h.schuler@pfaelzerwald.bv-pfalz.de
Sabine Hoos	Biosphärenreservat Pfälzerwald	Coordinatrice du projet LIFE Biocorridors en Allemagne	s.hoos@pfaelzerwald.bv-pfalz.de
Thomas Roussel	Biotope	Expertise des coléoptères saproxyliques	troussel@biotope.fr
Christophe Kimmel	DREAL Grand Est		christophe.kimmel@developpement-durable.gouv.fr
Stéphane Guillin	DREAL Grand Est/SEBP/PPPC	SINP Grand Est	stephane.guillin@developpement-durable.gouv.fr
Ernst Segatz	FAWF Trippstadt		ernst.segatz@wald-rlp.de
Patricia Balcar	FAWF Trippstadt		particia.balcar@wald-rlp.de
Johannes Becker	Forstamt Bienwald		johannes.becker@wald-rlp.de
Burkhard Steckel	Forstamt Johanniskreuz		burkhard.steckel@wald-rlp.de
Bodo Mahl	Forstamt Kaiserslautern		bodo.mahl@wald-rlp.de
Brice Noirot	Geopeka	Expertise cartographie	brice.noirot@geopeka.com
Christophe Rousson	Geopeka	Expertise cartographie	christophe.rousseau@geopeka.com
Markus Rink	Hirschkäferfreunde www.hirschkaefer-suche.de	Association pour l'étude et la protection du lucane	hirschkaefer-rink@t-online.de info@hirschkaefer-suche.de
Arnaud Horellou	INPN-SINP		arnaud.horellou@mnhn.fr
Mathieu Manceau	INPN-SINP		mathieu.manceau@mnhn.fr
Ludwig Simon	Landesamt für Umwelt		ludwig.simon@lfu.rlp.de
Joern Buse	Nationalpark Schwarzwald	Observateur scientifique, expertise des coléoptères saproxyliques	joern.buse@nlp.bwl.de
Ludovic Fuch	Office National des Forêts	Expertise coléoptères saproxyliques	
Bruno Mériguet	Office Pour les Insectes et leur Environnement	Etude du Lucane cerf-volant	bruno.meriguet@insectes.org
Loic Duchamp	PNR Vosges du Nord	RN des Rochers et tourbières du Pays de Bitche	l.duchamp@parc-vosges-nord.fr
Nadia Carmaux	PNR Vosges du Nord	Chargée de mission SIG	n.carmaux@parc-vosges-nord.fr
Olivier Laurent	PNR Vosges du Nord	Administration du projet LIFE Biocorridors	o.laurent@parc-vosges-nord.fr
Pishum Migraine	PNR Vosges du Nord	Chef de projet, coordinatrice du LIFE Biocorridors en France	p.migraine@parc-vosges-nord.fr
Sébastien Morelle	PNR Vosges du Nord	Chargé de mission N2000	s.morelle@parc-vosges-nord.fr

Nom	Organisation	Notes	Contact
Jean-Claude Génot	PNR Vosges du Nord SYCOPARC	Chargé de la protection de la nature	jc.genot@parc-vosges-nord.fr
Arno Thomaes	Research Institute for Nature and Forest (INBO)	Expertise modélisation pour le LCV	arno.thomaes@inbo.be
Henry Callot	Société Alsacienne d'Entomologie	Données de coléoptères	henry.callot@orange.fr
Christoph Best	Struktur- und Genehmigungsdirektion Süd	Comité technique projet LIFE Biocorridors	christoph.best@sdsued.rlp.de
Thomas Schlindwein	Struktur- und Genehmigungsdirektion Süd		thomas.schlindwein@sgdsued.rlp.de
Wolfgang Weigand	Struktur- und Genehmigungsdirektion Süd		wolfgang.weigand@sgdsued.rlp.de
Frank Köhler	www.colkat.de	Bases de données	frank.koehler@online.de
Michael Ley	Zentralstelle der Forstverwaltung	Cartes forestières Allemagne	michael.ley@wald-rlp.de
Günter Franz	Zentralstelle der Forstverwaltung	Souveraineté et services	guenter.franz@wald-rlp.de
Hartmut König	Zentralstelle der Forstverwaltung	Cartes forestières Allemagne	hartmut.koenig@wald-rlp.de
Jürgen Thielen	Zentralstelle der Forstverwaltung	Protection de la nature en forêt	juergen.thielen@wald-rlp.de
Thomas Sprengel	Zentralstelle der Forstverwaltung	Protection de la nature en forêt	thomas.sprengel@wald-rlp.de



Tableau 19. Réunions de travail

Date	Objet
15 mai 2017	Conseil scientifique du projet LCV (La Petite-Pierre, France)
Mai 2017	Échanges techniques et sur les méthodes d'analyse avec Arno Thomaes (Mantova, Italie)
20 novembre 2017	Présentation de résultats et discussion sur les acquisition de données (Strasbourg, France)
13 février 2018	Présentation des premiers résultats (Neustadt an der Weinstraße, Allemagne)
25 janvier 2019	Réunion de restitution de l'étude (Sturzelbronn, Allemagne)

## 7. Annexe 2. Carte détaillée des prédictions de Lucane

