

L'Institut Agro Rennes-Angers  
 Site d'Angers  Site de Rennes

<p>Année universitaire : 2021 - 2022 Spécialité : Mathématiques appliquées, statistiques Spécialisation (et option éventuelle) : Science des données pour la biologie</p>	<p><b>Mémoire de fin d'études</b></p> <p><input type="checkbox"/> d'ingénieur de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement)</p> <p><input type="checkbox"/> de master de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement)</p> <p><input type="checkbox"/> de l'Institut Agro Montpellier (étudiant arrivé en M2)</p> <p><input checked="" type="checkbox"/> d'un autre établissement (étudiant arrivé en M2)</p>
---	--

# Effets de mesures de conservation intégrative en forêt sur la dynamique d'apparition d'un dendromicrohabitat

Par : Pierre COTTAIS

*Soutenu à Rennes le 7 septembre 2022*

**Devant le jury composé de :**

Président :

Maître de stage : Fabien LAROCHE

Enseignant référent : François HUSSON

Autres membres du jury :

Sébastien LÉ, enseignant-chercheur

Benoit COURBAUD, enseignant-chercheur, co-encadrant

*Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle de l'Institut Agro Rennes-Angers*

Ce document est soumis aux conditions d'utilisation «Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France» disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>





**Effets de mesures de conservation  
intégrative en forêt sur la dynamique  
d'apparition d'un dendromicrohabitat**

Rédigé par : Pierre COTTAIS



# Remerciements

Je remercie tout d'abord la Fondation pour la recherche sur la biodiversité (FRB) pour le financement de mon stage, ainsi que l'Agence nationale de la recherche (ANR), sans qui le projet BloBiForM n'aurait pas été possible.

Je tiens à adresser toute ma gratitude à mon tuteur de stage Fabien LAROCHE, enseignant-chercheur et responsable du projet BloBiForM (ANR-19-CE32-0002-01) à l'UMR Dynafor<sup>1</sup> (INRAE Toulouse), qui, par sa disponibilité et la qualité de son encadrement, a su me conseiller et me guider durant l'intégralité de mon stage. Je remercie également mon co-encadrant Benoit COURBAUD, enseignant-chercheur au LESSEM<sup>2</sup> (INRAE Grenoble), qui m'a très bien accueilli à Grenoble pour me former sur son modèle de simulation *Samsara*, et qui a répondu à mes questions tout au long de mon stage, notamment sur ses travaux de recherche en modélisation de l'apparition de dendromicrohabitats. J'adresse également mes remerciements à Laurent LARRIEU, Nicolas GOUIX et Yvon GRZELEC, pour leur disponibilité et leur diligence dans les moments clés de ma mission, notamment par leur expertise respective.

J'adresse aussi mes remerciements à tous les membres de l'unité Dynafor, pour leur participation à l'ambiance de travail très agréable et bienveillante, et pour la richesse des discussions abordées, notamment lors de sessions comme les Retours du Jeudi organisées par Sylvie LADET, que je remercie particulièrement, ou bien lors des sessions du groupe de travail « statistiques des paysages » organisées par Fabien LAROCHE. Je remercie spécialement Agnes ARDANUY GABARRA, pour avoir participé volontairement à la seconde session d'acquisition de données en forêt de Grésigne.

Enfin, je tiens à remercier l'ensemble de l'équipe pédagogique du parcours Science des données de l'Institut Agro Rennes-Angers, pour leur pédagogie et grâce à qui j'ai eu l'opportunité d'enrichir mes connaissances scientifiques et mes compétences, ainsi que ma vision globale de l'enseignement.

---

1. Dynamiques et écologie des paysages AgriForestiers  
2. Laboratoire EcoSystèmes et Sociétés En Montagne



# Table des matières

<b>Table des illustrations</b>	<b>ii</b>
Figures . . . . .	ii
Tables . . . . .	ii
<b>Liste des acronymes</b>	<b>iii</b>
<b>Introduction</b>	<b>1</b>
Objectif et hypothèses . . . . .	2
<b>Matériel et méthodes</b>	<b>3</b>
Volet 1 : Calibration d'un modèle d'apparition de cavités basses . . . . .	3
Analogie avec l'analyse de survie . . . . .	3
Modèle hiérarchique de Weibull . . . . .	4
Données d'inventaire forestier . . . . .	5
Modèle statistique . . . . .	6
Estimation bayésienne . . . . .	7
Analyse de la qualité du modèle . . . . .	8
Volet 2 : Simulation de gestion sylvicole . . . . .	10
Modèle de simulation Samsara2 . . . . .	10
Définition de scénarios sylvicoles . . . . .	11
Simulation des scénarios avec Samsara2 . . . . .	11
<b>Résultats</b>	<b>12</b>
Qualité d'ajustement . . . . .	12
Distribution marginale <i>a posteriori</i> . . . . .	13
Qualité de prédiction . . . . .	15
Exploration des scénarios sylvicoles . . . . .	16
<b>Discussion</b>	<b>17</b>
<b>Références</b>	<b>21</b>

# Table des illustrations

## Figures

1	Taupin violacé . . . . .	2
2	Cavité à terreau de pied . . . . .	2
3	Modèle hiérarchique de Weibull avec effet placette aléatoire . . . . .	6
4	Nombre de chênes porteurs de CB par placette selon le modèle . . . . .	12
5	<i>Priors</i> et <i>posteriors</i> marginaux selon le modèle . . . . .	13
6	Corrélation des couples $(\alpha; \theta)$ et $(\alpha; \sigma_\delta)$ selon le modèle . . . . .	14
7	Probabilité qu'un chêne soit porteur d'une CB selon le modèle . . . . .	15
8	Nombre de chênes porteurs de CB par placette selon le modèle . . . . .	15
9	Densité moyenne de CB par placette selon le scénario . . . . .	17

## Tables

1	Lois <i>a priori</i> selon le modèle . . . . .	8
2	Acteurs ayant contribué aux scénarios de gestion sylvicole . . . . .	11

## Liste des acronymes

AF	Aménagement forestier
ANR	Agence nationale de la recherche
AUC	<i>Area under the curve</i>
CB	Cavité basse
CEN	Conservatoire des espaces naturels
CRPF	Centre régional de la propriété forestière
DMH	Dendromicrohabitat
DHP	Diamètre à hauteur de poitrine
DMA	Diamètre moyen d'apparition
GPS	<i>Global positioning system</i>
IC	Intervalle de crédibilité <sup>3</sup>
INRAE	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement
INPN	Inventaire national du patrimoine naturel
JAGS	<i>Just another gibbs sampler</i>
MCMC	<i>Markov chain Monte Carlo</i>
ONF	Office national des forêts
PSRF	<i>Potential scale reduction factor</i>
SIG	Système d'information géographique
TSF	Taillis-sous-futaie
UT	Unité territoriale

---

3. « IC » se réfère plus communément à l'intervalle de confiance mais ce dernier est écrit en toutes lettres dans ce mémoire afin de dédier l'acronyme « IC » à l'intervalle de crédibilité, plus fréquemment utilisé dans mon travail, notamment lors des évaluations de la qualité d'adéquation des modèles.



# Introduction

Ce stage s'inscrit dans un projet de l'ANR, porté par l'UMR Dynafor de l'INRAE, située au centre Occitanie-Toulouse. L'objectif de ce projet est d'explorer des orientations de gestion sylvicole pour favoriser la biodiversité au sein de forêts publiques gérées. Ce projet BloBiForM (*Block neutral models of Biodiversity to support Forest Management*) se focalise sur le cas particulier de la biodiversité saproxylique<sup>4</sup> en forêt de Grésigne, une forêt publique gérée ayant une activité de production de bois et possédant une biodiversité entomologique très riche.

La forêt domaniale de Grésigne est une chênaie située dans le département du Tarn. Sa superficie de 3 537 ha en fait la plus grande forêt de ce département. Elle est gérée par l'Office national des forêts (ONF), le gestionnaire des forêts publiques françaises. Cette ancienne forêt royale est caractérisée par une très grande richesse entomologique, cataloguée par Jean Rabil (1991), qui en fait la 3<sup>ème</sup> forêt d'Europe la plus riche en insectes, après les forêts de Fontainebleau en France et de Bialoweja en Pologne<sup>5</sup>. Le régime sylvicole de la Grésigne a longtemps consisté en une pratique de taillis sous futaie (TSF), jusqu'en 1875 où l'aménagement forestier fixe officiellement la transition du régime de TSF vers un régime de futaie régulière. L'ancien régime de TSF est un régime hybride, appliquant un régime de taillis, où les arbres, souvent multi-brins (c.-à-d. avec plusieurs tiges), sont issus de rejets sur souches. Ces arbres forment une strate jeune, dont la récolte est dédiée majoritairement au bois de chauffage, et dominée par une strate plus haute avec des arbres de plus gros diamètre issus de semis, dont le régime est celui de la futaie, d'où le nom de régime hybride de « taillis sous futaie ». Le régime de futaie régulière quant à lui, mobilise des interventions de sylviculture ayant pour objectif d'obtenir des peuplements de même classe d'âge. Les agents de l'ONF suivent ce régime sylvicole lors de leurs interventions au cours du cycle de développement des peuplements. À l'échelle de la forêt, cette transition du régime sylvicole de TSF vers celui de la futaie régulière est assez récente et, se faisant progressivement, elle est encore visible aujourd'hui dans certains peuplements d'arbres.

Un grand nombre d'espèces saproxyliques, notamment des coléoptères, vivent dans le bois en décomposition. Ces espèces jouent un rôle important dans le renouvellement de la matière ligneuse, dont elles sont l'un des principaux acteurs. Ce processus de saproxyllisation est fondamental dans le recyclage de la matière ligneuse et donc dans le cycle de vie global de la forêt (Kraus et Schuck, 2017). Les coléoptères saproxyliques vivent dans des habitats liés aux arbres vivants ou morts, regroupés sous le terme de dendromicrohabitats (DMH). Ces DMH peuvent prendre des formes diverses et variées. Le projet se focalise sur le Taupin violacé (voir FIGURE 1) qui vit dans un type de dendromicrohabitat spécifique : la cavité à terreau de pied, ou cavité basse de tronc (CB). La cavité basse est un dendromicrohabitat composé de bois mort en cours de décomposition (carie) et de terreau issu de l'humidification du sol. Le Taupin violacé, *Limoniscus violaceus* (P.W.J. Müller, 1821), est une espèce de coléoptère saproxylique menacée en Europe et dans le monde, voire en danger critique dans certaines régions de France<sup>6</sup>. La préservation de cette espèce, entre autres, est donc une priorité. La forêt de Grésigne est remarquable par une forte présence de cavités à terreau de pied, qui correspondent à un dendromicrohabitat

---

4. « Un organisme saproxylique dépend, pendant une partie de son cycle de vie, d'arbre moribond ou mort, debout ou à terre, ou de la présence d'autres organismes saproxyliques comme les champignons. » (site Internet du parc national des Pyrénées, consulté le 01/07/2022)

5. Site Internet de l'ONF, consulté le 27/07/2022

6. Site Internet de l'INPN, consulté le 02/08/2022

particulièrement favorable aux espèces saproxyliques (Ducasse et Brustel, 2008).



FIGURE 1 – Taupin violacé



FIGURE 2 – Cavité à terreau de pied

La mission de ce stage de master 2 est d'explorer l'implication de certaines pratiques sylvicoles sur la dynamique des cavités basses en forêt de Grésigne. On souhaite quantifier la dynamique d'apparition des cavités basses, afin de l'intégrer dans une gestion sylvicole de conservation de ces habitats favorables aux espèces saproxyliques comme le Taupin violacé, tout en maintenant de l'exploitation forestière. Cette stratégie de conservation s'ancre dans la notion d'approche « intégrative » de la gestion sylvicole. Il s'agit de prendre en compte, dans tous les peuplements exploités, des éléments favorables au maintien de la biodiversité dans le temps et dans l'espace avec, par exemple, un effort de rétention d'arbres porteurs de CB ou d'arbres ayant de grandes chances d'en développer (arbres à haut potentiel). Les scénarios de gestion à définir pour le cas particulier de la Grésigne s'inscrivent dans cette approche de gestion intégrative.

## Objectif et hypothèses

Dans toute l'étude, on se fonde sur une hypothèse qui consiste à faire l'analogie entre l'âge et le diamètre de l'arbre. On présuppose qu'il existe une relation croissante entre les deux, ce qui permet de modéliser la croissance d'un arbre par l'augmentation de son diamètre, et que cette relation est la même pour tous les arbres d'une espèce donnée et sur une placette d'observation donnée.

En plus de cette analogie présentée comme hypothèse de travail, deux hypothèses sont à tester pour pouvoir mener à bien cette mission. La première hypothèse postule que plus un arbre est vieux, plus la probabilité qu'il soit porteur d'une cavité basse est élevée (Vuidot *et al.*, 2011; Bouget *et al.*, 2014). Cela suppose l'existence d'un processus temporel de développement de la cavité basse. La seconde hypothèse postule que le processus de vieillissement de l'arbre, aussi appelé sénescence, augmente sa probabilité de développer une cavité basse. Autrement dit, la probabilité d'apparition d'une CB lorsque un jeune arbre croît est plus faible que celle d'un arbre déjà vieux qui n'a pas encore développé de cavité basse, à croissances radiales égales.

Le premier volet consiste à quantifier la dynamique d'apparition de la cavité basse en forêt de Grésigne, en s'appropriant un modèle calibré sur une base de données européenne d'arbres portant ou non des cavités basses. L'approche bayésienne permet de prendre en compte, dans l'estimation de paramètres, les connaissances antérieures sur un processus donné. Ici, cette approche semble être un choix judicieux pour intégrer l'information issue

de la base européenne dans les lois *a priori* du nouveau modèle pour la forêt de Grésigne. L'enjeu d'une telle approche est d'évaluer si la mobilisation des données européennes comme *a priori* améliore la qualité de prédiction du modèle. Il en résulte une comparaison d'un modèle avec un *a priori* informatif et d'un modèle sans information extérieure.

Une fois le modèle calibré sur la Grésigne, et donc la dynamique d'apparition de CB quantifiée, celle-ci doit être intégrée dans la gestion sylvicole pour pouvoir établir un compromis entre production de bois et conservation de la biodiversité. Le second volet a pour but de définir des scénarios de gestion intégrative basés sur l'effort de rétention des arbres porteurs de CB et à haut potentiel, afin d'en évaluer les effets à long terme sur la disponibilité dans le temps des cavités basses. On fait l'hypothèse qu'un effort de rétention lors des interventions de coupe de bois est favorable à la biodiversité associée. L'évaluation des effets des scénarios se fait par une simulation du développement d'un peuplement dans le temps. Le peuplement simulé va subir des interventions de coupe tout au long de sa croissance, ainsi que d'autres événements comme l'apparition de CB.

## Matériel et méthodes

### Volet 1 : Calibration d'un modèle d'apparition de cavités basses

Ce premier volet a pour objectif de calibrer un modèle d'apparition de cavités basses en Grésigne, à partir des données d'inventaire forestier, récoltées en forêt de Grésigne. Il est basé sur les travaux de Courbaud *et al.* (2017; 2022), présentant une modélisation de l'apparition du premier dendromicrohabitat comme une analyse de survie.

#### Analogie avec l'analyse de survie

L'analyse de survie modélise l'apparition d'un événement au cours du temps chez un individu. Dans le domaine médical, cela correspond par exemple au décès d'un patient atteint d'une maladie. La probabilité pour qu'un individu malade survive jusqu'à un instant  $t$  (c.-à-d. que le décès ne survienne pas avant cet instant  $t$ ) se modélise par la fonction de survie  $S(t) = \mathbb{P}(T > t)$ , avec  $T$  une variable aléatoire continue modélisant le moment du décès. Le raisonnement analogue peut être fait pour l'apparition de la première cavité basse (événement) au cours du temps sur un arbre (individu).

Cependant, déterminer le temps de croissance (c.-à-d. l'âge) d'un arbre nécessite un carottage aux coûts (physique et temporel) non négligeables lors des acquisitions de données sur le terrain. En effet, la tarière va endommager l'arbre observé et l'opération de carottage peut prendre un certain temps. Il est donc fait l'hypothèse d'une relation monotone croissante entre le diamètre à hauteur de poitrine (DHP) et le temps. Ce temps de croissance peut être représenté par le DHP de l'arbre, une variable plus facilement observable sur le terrain. On substitue donc le DHP au temps et l'accroissement en DHP à un pas de temps. La variable aléatoire continue  $D$  est définie comme étant le DHP d'un arbre sur lequel la CB apparaît. Il résulte de cette analogie les quatre fonctions suivantes, empruntées à l'analyse de survie.

1. La fonction de survie  $S(d)$  s'interprète alors comme la probabilité qu'un arbre ne porte toujours pas de CB lorsqu'il atteint le DHP  $d$ .

$$S(d) = \mathbb{P}(D > d) \tag{eq. 1}$$

2. Par complémentarité, la probabilité d'apparition d'une CB avant d'atteindre le

diamètre  $d$  s'écrit comme la fonction de répartition  $F(d)$ , de la variable aléatoire  $D$ . Elle s'interprète comme la probabilité qu'un arbre soit porteur d'une CB pour un DHP  $d$  donné.

$$F(d) = \mathbb{P}(D \leq d) = 1 - S(d) \quad (\text{eq. 2})$$

3. Si cette fonction de répartition  $F(d)$  est dérivable, on a alors la densité de probabilité  $f(d)$  associée à la variable aléatoire  $D$ .

$$f(d) = F'(d) \quad (\text{eq. 3})$$

4. Le taux  $h(d)$  d'apparition d'une CB lors d'un accroissement de diamètre  $\delta d$ , sachant que l'arbre n'en a pas encore au diamètre  $d$ , est appelé la fonction d'aléa (ou *hazard function* en anglais).

$$h(d) = \lim_{\delta d \rightarrow 0} \frac{\mathbb{P}(d \leq D \leq d + \delta d \mid D \geq d)}{\delta d} = \lim_{\delta d \rightarrow 0} \frac{\mathbb{P}((d \leq D \leq d + \delta d) \cap (D \geq d))}{\delta d \mathbb{P}(D \geq d)}$$

$$h(d) = \lim_{\delta d \rightarrow 0} \frac{\mathbb{P}(d \leq D \leq d + \delta d)}{\delta d \mathbb{P}(D \geq d)} = \frac{f(d)}{S(d)} \quad (\text{eq. 4})$$

Ce taux  $h(d)$  permet de rendre compte du processus de sénescence de l'arbre dans la probabilité d'apparition de la CB. En effet, il module cette hypothèse selon laquelle l'accroissement du DHP augmente de façon plus importante la probabilité de développer une CB pour un arbre déjà gros que pour un arbre de petit DHP. Un taux  $h(d)$  croissant correspondrait à cette hypothèse. Les travaux de Courbaud *et al.* (2022) montrent cependant que cette hypothèse n'est pas confirmée pour un certain nombre de DMH.

## Modèle hiérarchique de Weibull

Le modèle utilisé dans ce volet reprend le modèle hiérarchique de Weibull, développé par Courbaud *et al.* (2022) à partir de l'analogie avec l'analyse de survie présentée précédemment. Un modèle  $y$  est calibré pour chaque groupe de DMH. Cependant, l'écriture théorique de ces modèles reste identique. Le périmètre du projet BloBiForM se limitant à un seul type de DMH, la cavité basse, il est alors possible d'adapter l'écriture du modèle de Courbaud *et al.* (2022) aux données de la Grésigne.

La loi de Weibull à deux paramètres,  $\lambda$  (échelle) et  $k$  (forme) positifs, modélise les fonctions résultantes de l'analyse de survie :

— Fonction de survie :

$$S(d; \lambda, k) = e^{-\left(\frac{d}{\lambda}\right)^k} \quad (\text{eq. 1})$$

— Fonction de répartition de  $D$  :

$$F(d; \lambda, k) = 1 - e^{-\left(\frac{d}{\lambda}\right)^k} \quad (\text{eq. 2})$$

— Fonction de densité associée :

$$f(d; \lambda, k) = \frac{k}{\lambda} \left(\frac{d}{\lambda}\right)^{k-1} e^{-\left(\frac{d}{\lambda}\right)^k} \quad (\text{eq. 3})$$

— Fonction d'aléa (taux d'apparition de la première CB) :

$$h(d; \lambda, k) = \frac{k}{\lambda^k} d^{k-1} \quad (\text{eq. 4})$$

Le paramètre d'échelle  $\lambda$  est le diamètre auquel la probabilité de présence d'une CB vaut  $1 - e^{-1} \approx 0.63$ . Pour  $k$  fixé, si on augmente la valeur de  $\lambda$ , alors la plage du DHP où la première CB apparaît augmente. Cela inclut, entre autres, l'augmentation du DHP moyen d'apparition (DMA). Quant au paramètre de forme  $k$ , il module le processus de sénescence de l'arbre. Plus l'arbre vieillit, plus il a de chances de développer une CB. Si  $k$  est compris entre 0 et 1, le taux d'apparition de la première CB  $h(d)$  va décroître lorsque le DHP décroît. Cela correspond à un phénomène où plus l'arbre est jeune, plus il a de chance de développer une CB, ce qui est contraire à l'attendu biologique ici. À l'inverse, si  $k$  est supérieur à 1,  $h(d)$  va croître lorsque le DHP augmente. Cela correspond à l'hypothèse de la dynamique d'apparition des cavités basses, qui semble augmenter avec le vieillissement de l'arbre. Pour  $k = 1$ , on reconnaît la densité de la loi exponentielle « sans mémoire » (voir eq. 3). Dans ce cas, le taux d'apparition de CB reste constant malgré l'accroissement du DHP.

### Données d'inventaire forestier

Les données récoltées en Grésigne suivent un plan d'échantillonnage construit en neuf placettes circulaires d'un hectare. Elles ont été entièrement parcourues par les opérateurs terrain qui ont relevé de façon exhaustive les caractéristiques télémétriques et sylvicoles de toutes les tiges (ou brins) de plus de 17,5 cm de DHP. Pour chaque tige comptable<sup>7</sup>, les opérateurs ont relevé :

- la placette à laquelle elle appartient ;
- les coordonnées GPS<sup>8</sup> ;
- l'essence, ou espèce ;
- le diamètre à hauteur de poitrine ;
- et la présence ou l'absence de 26 types de DMH selon une nomenclature<sup>9</sup> pré-définie.

Ces neuf placettes « exhaustives » ont été parcourues lors d'une première session d'acquisition de données tout au long de l'année 2021, avant le début du présent stage. Une partie de ces données est exploitée comme jeu de calibration du modèle d'apparition de la première cavité basse.

Une seconde session d'acquisition de données, à laquelle j'ai participé en printemps 2022, se concentre uniquement sur le prélèvement de l'essence, ou espèce, du DHP et de la présence ou absence de cavité basse des arbres comptables. L'objectif est alors de construire un jeu de validation, afin de mesurer la qualité du modèle. Compte tenu de ce besoin, il n'est pas nécessaire de collecter les coordonnées GPS ou les autres types de dendromicrohabitats, ce qui du point de vue pratique, représente un gain de temps non négligeable sur le terrain. Lors de cette session d'acquisition, les arbres comptables ont été relevés dans sept nouvelles placettes.

---

7. Tige dont le DHP est supérieur ou égal à 17,5 cm

8. (EN) *Global Positioning System*

9. Cette nomenclature est une clé d'identification des types de dendromicrohabitats en les déclinant selon trois niveaux associés à leur stade de développement (Larrieu, 2019).

La phase d'exploration des données révèle une grande majorité de chênes, *Quercus*, ce qui est représentatif de la Grésigne, forêt à prédominance de chênes qualifiée de chênaie (plus de 80 % de chênes). Les essences restantes comme le charme et le hêtre, entre autres, sont bien moins nombreuses. À cause de cette différence considérable dans les effectifs, le nombre d'arbres porteurs de CB au sein de ces essences minoritaires est très faible. Il est donc décidé, pour le volet 1 de quantification de la dynamique d'apparition de CB, de ne travailler que sur les chênes.

### Modèle statistique

Un modèle « simple » serait de ne prendre en considération que le DHP comme variable explicative de l'apparition de la première CB. On peut cependant ajouter un effet aléatoire associé à la placette de mesure (voir description du protocole d'acquisition de données en partie précédente) pour prendre en compte la diversité des conditions de milieu et d'histoire des différents peuplements forestiers au sein de la Grésigne. On obtient alors le modèle hiérarchique décrit en FIGURE 3, qui attribue, pour chaque tige (ou brin)  $i$  d'une placette  $j$ , une probabilité  $p_{ij}$  de porter une CB. Cette probabilité est calculée à partir de la fonction de répartition de Weibull, où les paramètres  $\lambda_j$  et  $k$  sont définis de la façon suivante.

Le paramètre d'échelle  $\lambda_j$  dépend du paramètre  $\alpha$ , identique à tous les chênes du massif, et du paramètre  $\delta_j$ , correspondant à l'effet aléatoire placette.

$$\lambda_j = e^{\alpha + \delta_j} \quad (\text{eq. 5})$$

Le paramètre  $\delta_j$ , associé à l'effet aléatoire placette, suit une loi *a priori* définie selon un autre paramètre  $\sigma_\delta$  à estimer, appelé *hyperprior* (voir FIGURE 3).

Quant au paramètre de forme  $k$ , il est identique pour tous les chênes du massif, quelle que soit la placette dans laquelle la tige est observée.

$$k = e^\theta \quad (\text{eq. 6})$$

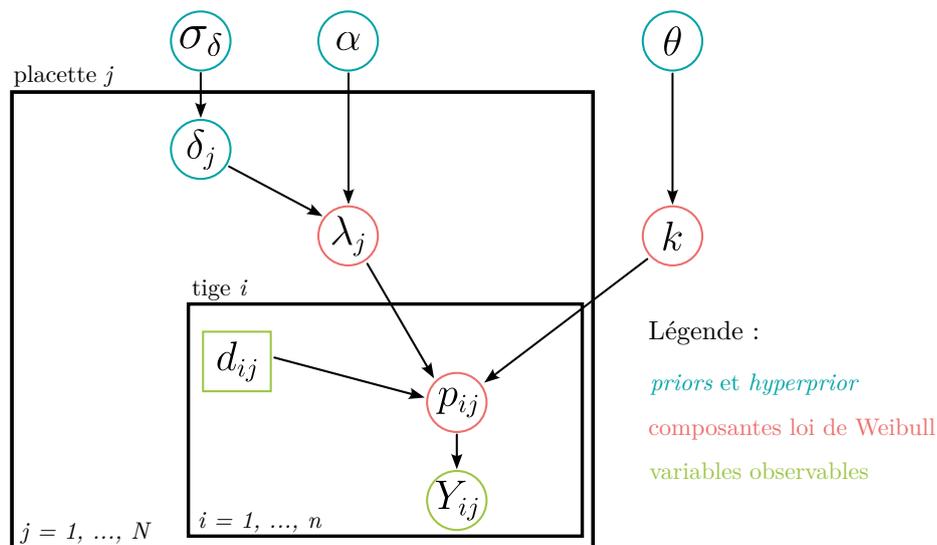


FIGURE 3 – Modèle hiérarchique de Weibull avec effet placette aléatoire

La FIGURE 3 reprend la notation usuelle des graphes orientés acycliques (*directed acyclic graphs* ou DAG en anglais). Les cercles représentent les paramètres ou variables aléatoires et le rectangle représente la variable déterministe  $d_{ij}$ , DHP de l'arbre  $i$  observé dans la placette  $j$ . La probabilité  $p_{ij}$  est calculée d'après la fonction de réparation  $F(d_{ij}; \lambda_j, k)$  est la probabilité qu'un chêne  $i$  porte une CB. La variable aléatoire booléenne  $Y_{ij}$  correspond justement à la présence ou absence de CB.

## Estimation bayésienne

Un des avantages de l'approche bayésienne dans l'estimation de paramètres est de pouvoir intégrer des connaissances extérieures dans les lois, ou distributions de probabilité, des paramètres *a priori*, aussi appelées *priors*<sup>10</sup>. Dans le cas du modèle hiérarchique de Weibull, à calibrer sur les données de la Grésigne, l'information extérieure provient des données utilisées dans les travaux de Courbaud *et al.* (2017; 2022). Ces données regroupent plus 80 000 arbres de 19 espèces différentes à travers des jeux de données de forêts en Europe et Iran, justifiant ainsi l'appellation volontairement simplifiée de base de données « européenne ». Ces forêts sont gérées ou non gérées et peuvent être divisées en plusieurs sites. Dans leur modélisation de l'apparition de DMH, Courbaud *et al.* ont agrégés les types de DMH en 11 groupes, où un modèle est calibré pour chaque groupe, dont celui des CB.

Grâce à l'exploitation de la base européenne dans la modélisation, il est maintenant possible de mobiliser cette information extérieure dans les *priors* du modèle d'apparition de CB à calibrer sur la Grésigne. Plus précisément, les *priors* sont construits à partir des distributions *a posteriori* du modèle de Courbaud *et al.*. On obtient donc un modèle dont les *priors* des paramètres  $\alpha$  et  $\theta$  sont qualifiés d'« informatifs ».

En revanche, dans la version de la calibration où cette information extérieure n'est pas mobilisée, les *priors* sont qualifiés de non informatifs. Ces lois *a priori* sont définies sur des gammes de valeurs très larges et centrées sur 0 (voir TABLE 1 ci-dessous).

Le modèle calibré sur les données européennes n'intègre pas d'effet aléatoire placette à l'intérieur des massifs forestiers recensés. Ainsi, dans les deux types de calibration envisagés ici (*priors* informatifs ou non informatifs), le paramètre  $\sigma_\delta$ , associé à cet effet, suit une loi *a priori* non-informative.

Les *priors* informatifs  $\alpha$  et  $\theta$ , présentés en TABLE 1, sont construits à partir du vecteur d'espérances  $M$  et de la matrice de variance-covariance  $\Sigma$ , issues des distributions *a posteriori* du modèle calibré sur les données de la base européenne par Courbaud *et al.* (2022) (voir équations ci-après). Dans cet ancien modèle, les paramètres  $\alpha'$ ,  $\beta$ , et  $\gamma$  correspondent respectivement aux effets espèce, gestion et site associés au paramètre d'échelle  $\lambda'$  de la loi de Weibull. Quant à  $\theta'$ , représentant également l'effet espèce, il est associé au paramètre de forme  $k'$ . L'effet site est propre aux données européennes. Celles-ci sont structurées telles que les individus (arbres) appartient à un site, lui-même appartenant à une forêt, laquelle est gérée ou non. Ainsi, le *prior* de  $\alpha$  du nouveau modèle intègre les effets espèce  $\alpha'$  et gestion  $\beta$  (forêt gérée) ainsi que la variance de l'effet site  $\sigma_\gamma$ . Le *prior* de  $\theta$  intègre quant à lui uniquement l'effet espèce  $\theta'$  (voir TABLE 1).

10. L'anglicisme lexical de *prior* est utilisé pour simplifier la lecture, notamment celle des graphiques.

$$M = \begin{bmatrix} \mu_{\alpha'} \\ \mu_{\beta} \\ \mu_{\gamma} \\ \mu_{\theta'} \end{bmatrix} \quad \text{et} \quad \Sigma = \begin{bmatrix} \sigma_{\alpha'}^2 & \text{cov}(\alpha, \beta) & \text{cov}(\alpha, \gamma) & \text{cov}(\alpha', \theta') \\ \text{cov}(\beta, \alpha') & \sigma_{\beta}^2 & \text{cov}(\beta, \gamma) & \text{cov}(\beta, \theta') \\ \text{cov}(\gamma, \alpha') & \text{cov}(\gamma, \beta) & \sigma_{\gamma}^2 & \text{cov}(\sigma_{\gamma}, \theta') \\ \text{cov}(\theta', \alpha') & \text{cov}(\theta', \beta) & \text{cov}(\theta', \gamma) & \sigma_{\theta'}^2 \end{bmatrix}$$

$$\rho = \begin{bmatrix} \rho_{\alpha'} \\ \rho_{\beta} \\ \rho_{\gamma} \\ \rho_{\theta'} \end{bmatrix} \quad \text{tel que} \quad \rho \sim \mathcal{N}(M; \Sigma)$$

Lois du modèle avec <i>priors</i> informatifs	Lois du modèle avec <i>priors</i> non informatifs
$\alpha = \rho_{\alpha'} + \rho_{\beta} + \rho_{\gamma}$	$\alpha \sim \mathcal{N}(0; 100)$
$\theta = \rho_{\theta'}$	$\theta \sim \mathcal{N}(0; 100)$
$\delta \sim \mathcal{N}(0; \sigma_{\delta})$	$\delta \sim \mathcal{N}(0; \sigma_{\delta})$
$\sigma_{\delta} \sim \mathcal{U}_{[0;10]}$	$\sigma_{\delta} \sim \mathcal{U}_{[0;200]}$

TABLE 1 – Lois *a priori* selon le modèle

L'estimation des distributions *a posteriori* des paramètres du modèle calibré sur la Grésigne se fait par une méthode d'échantillonnage de Monte-Carlo par chaînes de Markov (MCMC). L'estimation est calculée sous JAGS (*Just Another Gibbs Sampler*), en paramétrant un algorithme d'échantillonnage de Gibbs. Cependant, la paramétrisation en amont et le traitement des résultats de l'estimation se font sous le langage de programmation R. Le *package* `runjags` (Denwood, 2016, v2.2.1-7) fait le lien entre R et JAGS.

## Analyse de la qualité du modèle

**Convergence des estimations** Le premier point d'analyse de la qualité du modèle est de vérifier que la distribution *a posteriori* des paramètres estimés provient bien d'une distribution stationnaire. Autrement dit, il s'agit de vérifier que les chaînes MCMC convergent bien. En plus de la visualisation graphique de la convergence des chaînes, le facteur de réduction d'échelle potentiel (PSRF) peut être utilisé comme un indicateur de cette convergence (Brooks et Gelman, 1998). Pour chaque distribution *a posteriori*, plus le PSRF est proche de 1, plus la convergence des chaînes est jugée bonne. Si le PSRF est supérieur au seuil conventionnel de 1.1, on considère que les chaînes n'ont pas convergé. Dans ce cas, plusieurs solutions sont possibles pour améliorer la convergence : ne conserver qu'une partie finale des chaînes, augmenter la valeur du paramètre de « rodage » (ou *burn in* en anglais) en entrée de l'algorithme ou bien reparamétriser les *priors* du modèle.

Une fois la convergence confirmée, il faut également vérifier que les valeurs estimées des paramètres peuvent être considérées comme des tirages indépendants dans la distribution *a posteriori*. Autrement dit, il s'agit de vérifier qu'il n'y a pas d'autocorrélation entre des valeurs successives échantillonnées dans les chaînes, potentiellement causée par le processus numérique de l'échantillonneur de Gibbs. La visualisation de l'autocorrélation par un graphique de décalage emprunté à l'analyse de séries temporelles est, là aussi, très

efficace pour mettre en évidence l'indépendance des tirages dans la distribution *a posteriori*. De plus, le nombre de valeurs efficaces (valeurs considérées comme indépendantes) peut être mis en comparaison avec le nombre total de valeurs estimées. S'il y a peu de valeurs efficaces, relativement au nombre total de valeurs estimées, alors il y a une forte autocorrélation entre les valeurs. Dans ce cas, il faut jouer sur les paramètres d'entrée de l'algorithme, comme la longueur des chaînes ou l'élagage (ou *thinning* en anglais, c.-à-d. ne conserver qu'une partie des valeurs) ou bien reparamétriser les *priors* du modèle pour réduire l'autocorrélation.

**Qualité d'ajustement** Dès lors que ces deux aspects inhérents à la convergence du modèle sont vérifiés, il est alors envisageable d'étudier la qualité d'ajustement du modèle avec *priors* informatifs ou non informatifs, à partir du jeu de données de calibration et des jeux de données simulés par une approche de rééchantillonnage paramétrique, c.-à-d. à partir du modèle estimé et des données observées (ou méthode *bootstrap* en anglais). Pour ce faire, on considère plusieurs statistiques : le nombre total de chênes porteurs de CB, le nombre total de chênes porteurs de CB au sein de chaque placette et la vraisemblance des jeux de paramètres  $(\alpha; \theta; \delta_j)$  estimés, ou plus exactement la log-vraisemblance. Pour chaque statistique, la qualité d'ajustement est évaluée de la façon suivante :

- tirage des 8 000 jeux de paramètres estimés dans la distribution jointe *a posteriori* sous forme d'autant de vecteurs de paramètres  $(\alpha; \theta; \delta_j)$ , pour ensuite calculer les couples de paramètres  $\lambda_j$  et  $k$  de la fonction de répartition de Weibull (voir eq. 5 et eq. 6 page 6) ;
- calcul de 8 000 probabilités  $p_i$  d'apparition de CB, pour le DHP de chaque tige  $i$  des 1 463 tiges du jeu de données utilisé pour la calibration du modèle (voir eq. 2 page 4) ;
- tirages aléatoires dans la loi de Bernoulli pour chaque probabilité  $p_i$ , afin d'attribuer ou non une CB à chaque tige  $i$  ;
- calcul de la statistique sur les 1 463 tiges.

On obtient ainsi une distribution empirique de 8 000 valeurs simulées de la statistique par *bootstrap* à partir du modèle, à comparer avec la statistique calculée sur les valeurs réellement observées.

**Qualité de prédiction** Enfin, le choix du type de *priors* du modèle est également motivé par la performance prédictive de celui-ci. Les statistiques utilisées dans cette étape sont identiques à celles de l'analyse de la qualité d'ajustement, à ceci près que les données utilisées sont celles du jeu de validation, récoltées lors de la seconde session d'acquisition en avril 2022. La distribution empirique de chaque statistique est comparée avec la valeur observée de cette statistique. On parle ici de prédiction car l'approche par échantillonnage paramétrique se fait sur les données de validation, qui n'ont pas été utilisées pour calibrer le modèle et donc pour l'estimation des paramètres. En plus de ces trois statistiques, l'AUC semble, au premier abord, un indicateur, sinon pertinent, du moins très utilisé dans la comparaison de la performance prédictive entre modèles.

## Volet 2 : Simulation de gestion sylvicole

Ce second volet consiste à évaluer l'effet de scénarios sylvicoles intégrant un effort de rétention plus ou moins ambitieux, en matière de préservation, selon le scénario. Cette évaluation exploratoire se fait à l'aide d'un simulateur de croissance de peuplement à l'échelle de l'arbre.

### Modèle de simulation Samsara2

Samsara2 est un modèle spatialement explicite de dynamique forestière à l'échelle d'un individu (arbre) au sein d'un peuplement (Courbaud *et al.*, 2015). Ce modèle, entre autres, est implémenté sur la plateforme de simulation de dynamique forestière Capsis<sup>11</sup> (Dufour-Kowalski *et al.*, 2012). Le modèle Samsara2 simule le cycle de vie d'un individu (allant du semis à l'arbre mort) en prenant en compte plusieurs éléments, intrinsèques aux caractéristiques de l'arbre (essence, DHP) mais également relatifs au peuplement dans lequel il évolue (coordonnées spatiales, proximité avec les autres arbres, éclaircissement au sol, *etc.*). La croissance d'un individu est modélisée par un processus stochastique dépendant de l'éclaircissement de l'arbre, c.-à-d. de la quantité de lumière interceptée par le houppier. L'incidence de la lumière et la quantité de lumière interceptée sont influencées par la taille de l'arbre, mais également celle de ses voisins. Le modèle intègre donc la notion de compétition pour la lumière entre les arbres d'un même peuplement. On peut évaluer sa croissance par son augmentation en surface terrière dans le temps. Ce processus stochastique de croissance d'arbre intègre le modèle d'apparition de CB calibré dans le volet 1 (voir Modèle statistique page 6), en récupérant la distribution médiane *a posteriori* estimée par ce modèle. Le modèle de croissance d'arbre est construit à partir de la fonction d'aléa  $h(d)$  (voir Analogie avec l'analyse de survie page 3), déterminante pour la simulation de l'apparition des CB dans le peuplement virtuel de Samsara2.

Dans le logiciel de simulation Capsis4, le peuplement est représenté par un polygone défini par les bords d'une placette. C'est au sein de ce polygone que se développent les arbres. Au cours du temps, le peuplement peut subir diverses interventions sylvicoles, suivant la paramétrisation de la simulation, c.-à-d. selon le scénario que l'on souhaite simuler. Un scénario est paramétré par un ensemble d'événements simulant un itinéraire sylvicole. On y définit par exemple des seuils minimaux de surface terrière<sup>12</sup> du peuplement après coupe (en m<sup>2</sup> par ha) et de nombre d'arbres de rétention par ha.

Une simulation représente l'application d'un scénario sylvicole pendant une période donnée. Il faut d'abord définir une initialisation du peuplement, correspondant à la liste de tous les arbres observés dans une placette. Une fois le peuplement initialisé, celui-ci va subir autant de séries d'interventions sylvicoles qu'il y a de scénarios. Ainsi, un scénario de gestion correspond à une série d'interventions. Bien que les simulations (c.-à-d. scénarios), soient indépendantes les unes des autres, elles partent toutes de la même initialisation (c.-à-d. peuplement). Cela permet, une fois les simulations terminées, de pouvoir comparer les scénarios entre eux.

---

11. CAPSIS : Croissance d'Arbres en Peuplement avec Simulation d'Itinéraires Sylvicoles

12. La surface terrière d'un arbre correspond à la surface de la section transversale de sa tige à hauteur de poitrine. La surface terrière d'un peuplement est donc la somme des surfaces terrières des arbres qui le constituent.

## Définition de scénarios sylvicoles

Afin de construire des scénarios pertinents, plusieurs acteurs ont été rencontrés pour discuter du choix de scénarios à simuler, de leur pertinence au regard de la gestion par approche intégrative ou bien simplement de leur faisabilité sur le terrain (voir TABLE 2 ci-dessous).

Nom	Fonction	Contribution
Nicolas GOUIX	Chef de projets au CEN Occitanie	Aménagement forestier et SIG
Yvon GRZELEC	Responsable de l'UT Grand Ouest à l'ONF	Itinéraires d'interventions
Laurent LARRIEU	Technicien CRPF Occitanie et chercheur écologue	Expertise CB et sylviculture

TABLE 2 – Acteurs ayant contribué aux scénarios de gestion sylvicole

Il résulte de ces discussions trois scénarios. Ces trois scénarios ont en commun leur durée dans le temps, établie à 285 ans. Cette durée est volontairement supérieure à l'âge limite d'exploitabilité d'un arbre (250 ans) afin d'intégrer un cycle complet de sylviculture opérée sur un peuplement. Le premier scénario peut être vu comme un scénario « témoin » servant de référence. Il reflète une croissance de peuplement en libre évolution, sans aucune intervention, que ce soit des coupes d'éclaircie ou de récolte de bois. Les deux autres scénarios se focalisent sur les effets de l'effort de rétention envers les arbres porteurs de CB. Pour mesurer ces effets, le seul paramètre variant entre ces deux scénarios est le nombre d'arbres de rétention. L'un des scénarios comporte un seuil d'au moins 3 arbres de rétention par hectare, il est appelé *business as usual* (BAU) car il correspond à la gestion habituelle opérée par l'ONF en forêt de Grésigne, avant l'adoption de la charte du réseau Natura 2000, en 2007. L'autre scénario élève légèrement ce seuil à 10 arbres de rétention par hectare, équivalent à environ 15 m<sup>3</sup> par ha, comme le préconisait le document d'objectifs du site Natura 2000 de la Grésigne (2005).

## Simulation des scénarios avec Samsara2

Ces trois scénarios sont appliqués sur les six placettes du jeu de données de calibration qui sont situées dans la série n°1 de l'aménagement forestier (AF 2004-2021, ONF), zone de production de bois. La définition des interventions de coupes, relatives aux deux scénarios de production de bois (BAU et effort de rétention), s'inspire des interventions prévues dans le document d'aménagement forestier de l'ONF. L'ordre d'application de ces interventions et la quantité de bois récolté à chacune d'entre elles diffèrent selon l'âge du peuplement, ou plus précisément selon son stade de développement. Les deux scénarios de production de bois sont adaptés à chaque placette, tout en maintenant le même nombre d'arbres de rétention pour toutes les interventions de coupes d'un même scénario. Le choix des arbres de rétention priorise les arbres porteurs de CB, puis les gros arbres à haut potentiel dans un second temps. Pour les deux scénarios impliquant des interventions, les coupes d'éclaircie interviennent tous les 10 ans, sauf pour les jeunes peuplements au stade du perchis (tiges fines), où les coupes d'éclaircie sont espacées de 15 ans. Le cycle complet de sylviculture se clôt avec une coupe définitive de régénération du peuplement, alors âgé de 250 ans.

La croissance d'un arbre étant modélisée par un processus stochastique dans le modèle Samsara2, chaque simulation va être différente pour un même scénario. Afin d'être plus confiant dans l'interprétation des résultats de simulation, 30 répétitions des trois scénarios simulés (BAU, effort de rétention et libre évolution) sont effectués, pour chacune des six placettes.

# Résultats

## Qualité d'ajustement

Cette première section de présentation des résultats reprend la méthodologie d'analyse de la qualité du modèle avec *priors* informatifs issus de la base européenne et du modèle sans information extérieure, ou avec *priors* non informatifs.

La convergence des estimations est améliorée par l'augmentation des paramètres d'entrée de l'algorithme : la longueur des quatre chaînes MCMC, le *burn in* et de la taille des échantillons tirés dans les chaînes (équivalent au *thinning*). Le modèle calibré, dont les résultats sont présentés ci-après, est paramétré avec 4 chaînes MCMC de longueur  $10^7$ , avec un *burn in* de  $10^4$  et un pas d'élagage de 5 000. Grâce à ce dernier paramètre, l'autocorrélation entre les valeurs échantillonnées est nettement réduite (entre  $-0.05$  et  $0.05$ ). L'amélioration de la convergence entraîne néanmoins une augmentation significative du temps d'exécution de l'algorithme passant de quelques minutes à plusieurs heures.

D'après la première statistique (nombre total de chênes porteurs), calculée à partir des données simulées par le modèle et des données observées, la qualité d'ajustement du modèle semble bonne, quel que soit le type de *priors* du modèle. Les valeurs issues des données simulées sont bien distribuées autour du nombre total de chênes porteurs effectivement observé.

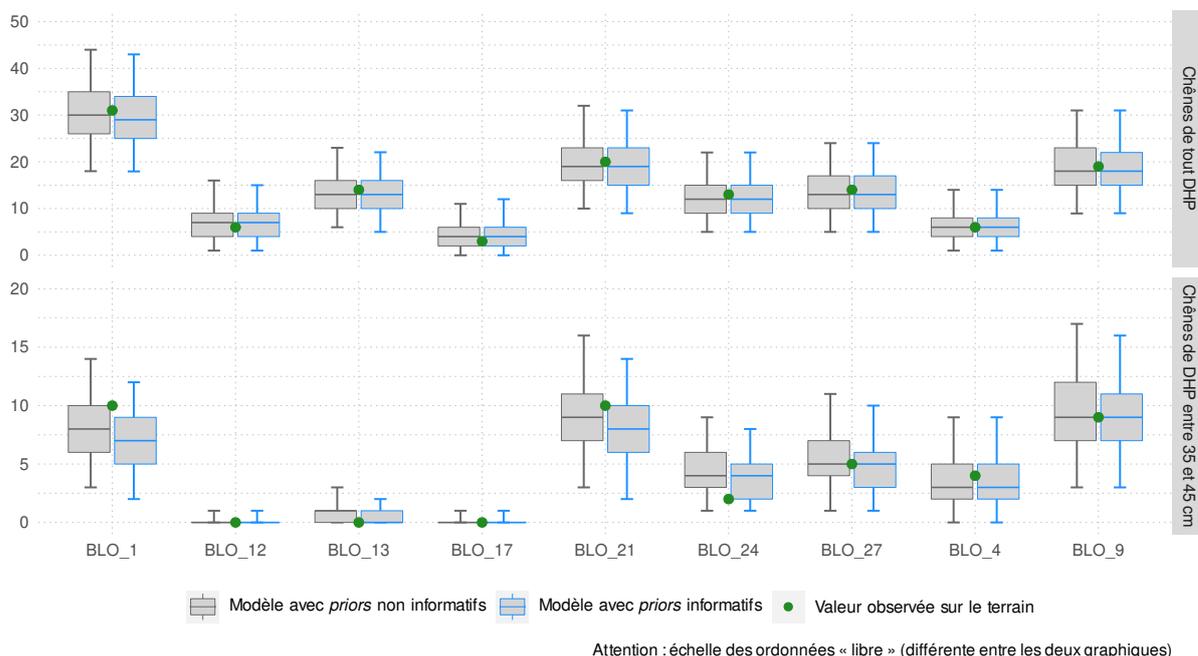


FIGURE 4 – Nombre de chênes porteurs de CB par placette selon le modèle

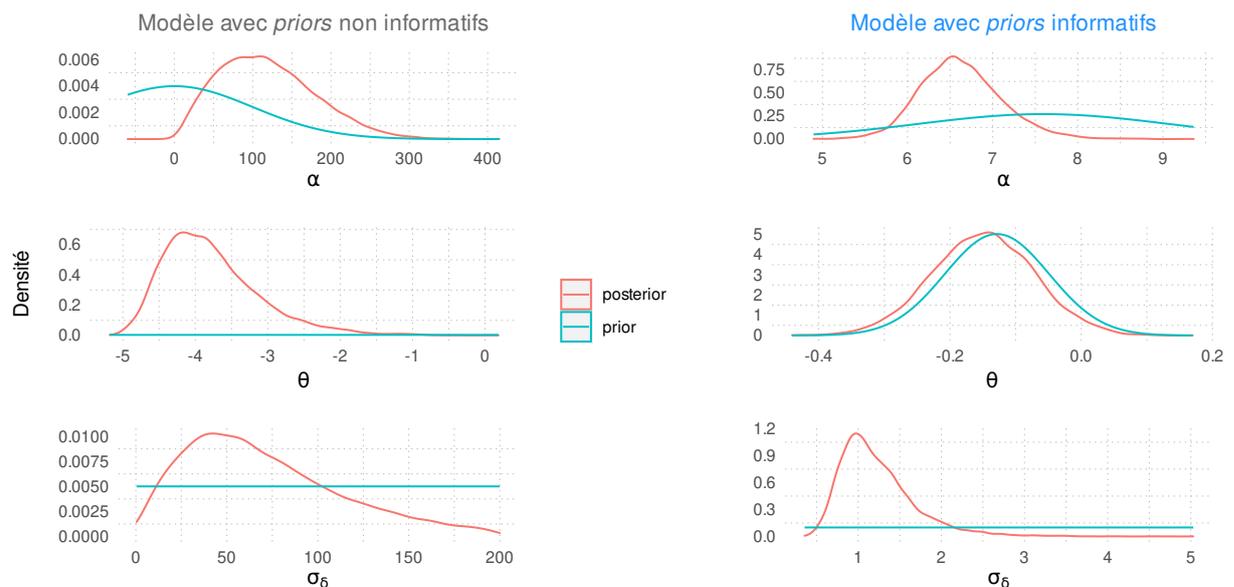
La FIGURE 4 présente la distribution du nombre de chênes porteurs de CB par placette simulée selon le modèle avec *priors* informatifs ou non informatifs. D'un point de vue général, l'estimation du nombre de chênes porteurs de CB, tous DHP confondus, se comporte de façon identique pour les deux types de modèle. La qualité d'ajustement semble plutôt bonne (valeur observée comprise dans les intervalles interquartiles des valeurs simulées) au regard de cette statistique.

Lorsque l'on étudie la qualité de l'ajustement des deux types de modèle pour une classe de DHP entre 35 et 45 cm, la qualité d'ajustement semble un peu moins bonne. On remarque un léger effet de sous-estimation dans les placettes BLO\_1 et BLO\_21 et un léger effet de surestimation dans la placette BLO\_24. De plus, l'estimation ne se comporte pas tout à fait de la même façon selon le type de *priors*. En effet, pour ces trois placettes où les valeurs simulées sont éloignées de la valeur observée (valeur observée en dehors de l'intervalle interquartile), le modèle avec *priors* informatifs semble généralement prédire moins de chênes porteurs de CB que celui avec *priors* non informatifs. Cette tendance à prédire moins de CB pour le modèle avec *priors* informatifs provoque une meilleure qualité d'ajustement de celui-ci dans des placettes où le nombre de chênes porteurs de CB est surestimé par ces modèles. Ces phénomènes de sous-estimation et surestimation sont qualifiés de « légers » car la valeur observée de cette statistique est toujours comprise dans l'intervalle de crédibilité à 95% (IC correspondant à 95% des valeurs simulées et représenté par les moustaches des boîtes sur le graphique FIGURE 4).

## Distribution marginale *a posteriori*

La méthode MCMC renvoie une distribution *a posteriori* jointe des paramètres  $(\alpha; \theta; \sigma_\delta)$  estimés. L'effet placette étant aléatoire, on étudie plutôt le paramètre  $\sigma_\delta$ , qui correspond à l'écart-type des estimations du paramètre  $\delta_j$ . Il est important de noter que dans le modèle, le *prior* de ce paramètre est non informatif (loi uniforme  $\mathcal{U}_{[0;10]}$  pour le modèle avec *priors* informatifs). En effet, il n'existe pas, dans la base européenne, d'effet placette. Les arbres sont situés dans des sites au sein d'une forêt donnée, certaines forêt ne possédant parfois qu'un seul site, mais il n'y a pas de granularité plus fine avec un découpage de site en plusieurs placettes.

On adopte deux niveaux d'analyse sur la visualisation des distributions marginales : la comparaison entre le *prior* et le *posterior* d'un paramètre au sein d'un même type de modèle et la comparaison entre les *posteriors* d'un paramètre des deux types de modèle.



Attention : échelle des abscisses « libre » (différente entre les graphiques)

FIGURE 5 – *Priors* et *posteriors* marginaux selon le modèle

D'après les abscisses de la FIGURE 5, les *posteriors* des trois paramètres  $\alpha$ ,  $\theta$  et  $\sigma_\delta$  du modèle avec *priors* non informatifs sont distribués sur des plages de valeurs très larges, par rapport au modèle avec *priors* informatifs. Ces grandes plages de valeurs en abscisse révèlent une forte variance des *posteriors* et donc une difficulté d'identifiabilité des paramètres du modèle. Comme on l'a vu plus haut, cette mauvaise qualité d'estimation n'implique pas une mauvaise qualité d'ajustement. En effet, ces trois paramètres co-varient fortement (voir corrélation des couples  $(\alpha; \theta)$  et  $(\alpha; \sigma_\delta)$  en FIGURE 6). Autrement dit, les fortes valeurs estimées d'un même jeu de paramètres se compensent dans l'expression de la fonction de répartition de Weibull (eq. 2), offrant ainsi un grand nombre de valeurs candidates pour ce jeu de paramètres, et donc de modèles paramétriques possibles, mais donnant les mêmes probabilités de présence ou absence de CB.

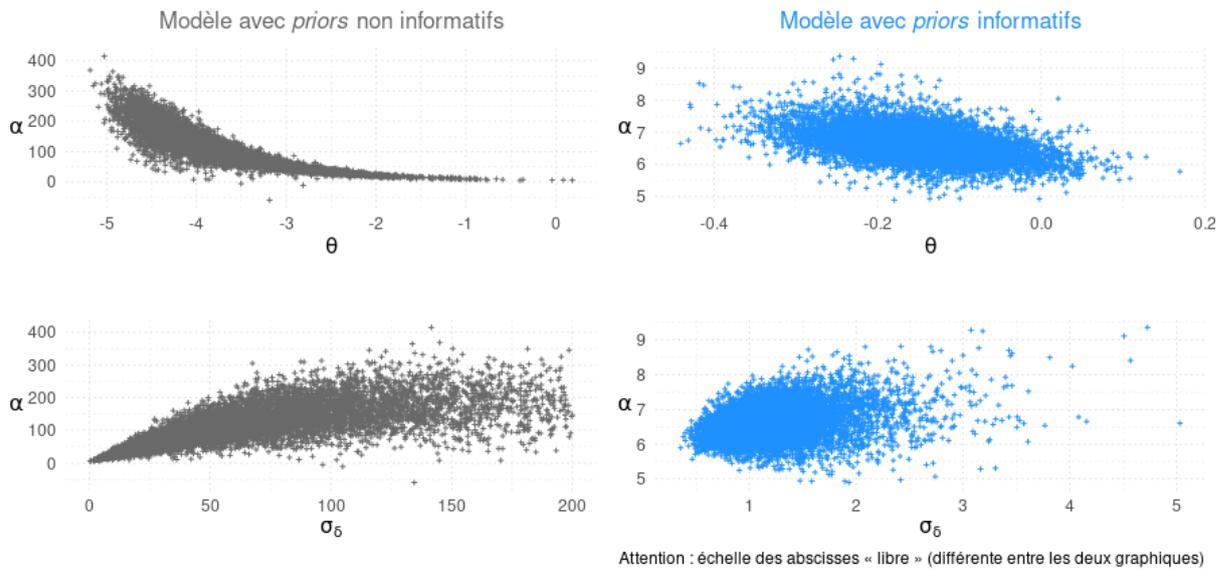


FIGURE 6 – Corrélation des couples  $(\alpha; \theta)$  et  $(\alpha; \sigma_\delta)$  selon le modèle

D'après la FIGURE 6, le modèle avec *priors* non informatifs illustre bien cette co-variance dans l'estimation des trois paramètres  $\alpha$ ,  $\theta$  et  $\sigma_\delta$ . Pour les couples  $(\alpha; \theta)$  et  $(\alpha; \sigma_\delta)$  du modèle avec *priors* non informatifs, le coefficient de corrélation<sup>13</sup> vaut respectivement  $-0.85$  et  $0.70$ . Tandis que pour le modèle avec *priors* informatifs, ce coefficient vaut respectivement  $-0.47$  et  $0.31$ .

Les paramètres du modèle « informé »<sup>14</sup> sont estimés sur des plages de valeurs bien plus petites que le modèle « non informé ». De plus, les courbes des *posteriors* de  $\alpha$  et  $\sigma_\delta$  sont plus resserrées et en forme de cloche. Le paramètre  $\theta$ , estimé autour de  $-0.15$ , est négatif. Le paramètre de forme  $k$  est donc inférieur à 1, ce qui correspond à un taux d'apparition de CB décroissant avec le DHP (voir Modèle hiérarchique de Weibull page 4). Néanmoins,  $\theta$  est également négatif dans le modèle non informé, et ce en dépit de ses larges variations (entre  $-5$  et  $-2$ ).

13. Le coefficient de corrélation calculé ici est le coefficient de corrélation linéaire de Pearson.

14. Les expressions de modèle informé ou non informé ont pour but de simplifier la lecture, au détriment des expressions respectives de modèle à *priors* informatifs ou non informatifs, plus rigoureuses.

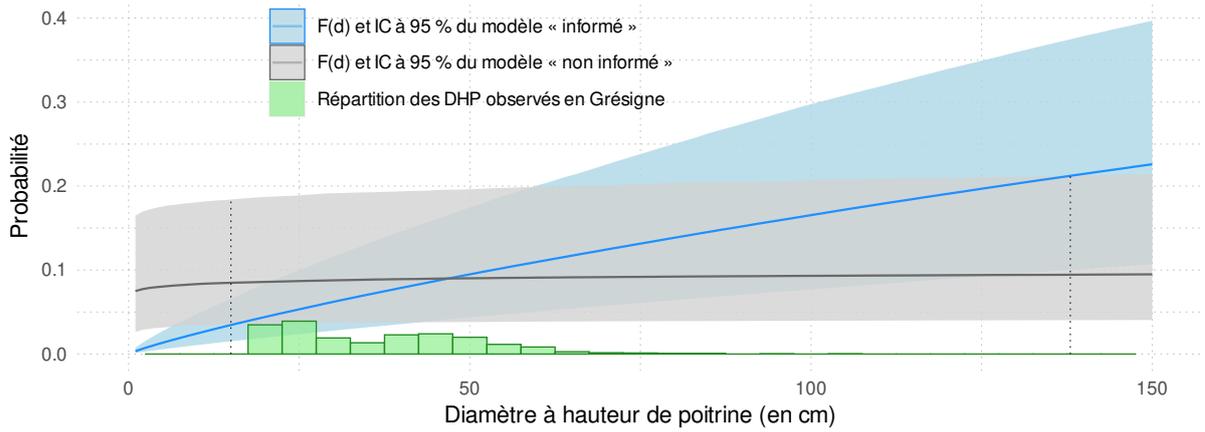


FIGURE 7 – Probabilité qu’un chêne soit porteur d’une CB selon le modèle

La FIGURE 7, illustrant les fonctions de répartition selon le type de *priors*, montre un autre effet de l’introduction d’information *a priori* dans la calibration : la croissance de la courbe grise (modèle avec *priors* non informatifs), est très faible, ce qui signifie que cette probabilité change peu avec le diamètre du chêne. À l’inverse, la croissance de la courbe bleue est plus forte, ce qui signifie que la probabilité qu’un chêne soit porteur d’une CB est de plus en plus élevée que son diamètre est grand. Cependant, on remarque que l’enveloppe (IC à 95 %) associée au modèle non informé chevauche la courbe « médiane » du modèle informé, pour l’ensemble des DHP observables (entre 15 et 140 cm). La probabilité qu’un chêne soit porteur d’une CB est donc globalement similaire pour les deux types de *priors*, du moins pour des valeurs de DHP observables sur le terrain.

## Qualité de prédiction

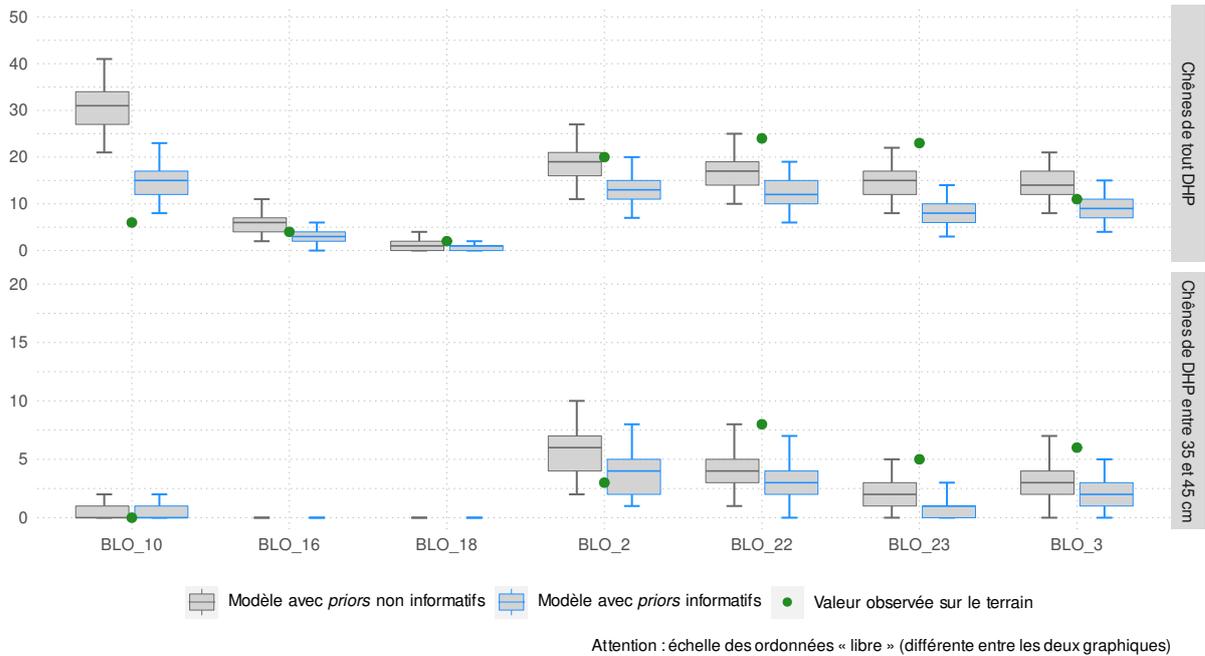


FIGURE 8 – Nombre de chênes porteurs de CB par placette selon le modèle

D'après la FIGURE 8, la qualité de prédiction ne semble pas très bonne, en particulier pour le modèle avec *priors* informatifs. La valeur observée du nombre de chênes porteurs de CB se situe parfois en dehors de l'intervalle interquartile, voire de l'IC à 95 %. Quant à la comparaison de performance prédictive, on retrouve le même résultat qu'en analyse de la qualité d'ajustement, où le nombre de chênes porteurs de CB prédit par le modèle informé est toujours plus faible que celui prédit par le modèle non informé. Dans la plupart des placettes, cette statistique est sous-estimée, ce qui rend le modèle non informé globalement meilleur, sauf pour la placette BLO\_10 où le nombre de chênes porteurs de CB est surestimé. En résumé, le modèle avec *priors* informatifs prédit encore moins de chênes porteurs de CB que le modèle avec *priors* non informatifs.

L'AUC est identique pour les deux types de *priors* du modèle ( $AUC = 0.63$ ). Cela s'explique par le fait que les probabilités prédites par le modèle (eq. 2) sont toujours ordonnées de la même façon que l'ordre des DHP, quel que soit le type de *priors*. La construction des seuils de la courbe ROC<sup>15</sup> n'étant sensible qu'à l'ordre des probabilités, la courbe est identique pour les deux types d'*a priori*. Dans ce cas, où l'on souhaite comparer la performance prédictive, l'AUC n'est pas un indicateur pertinent. Un point d'amélioration de l'analyse de la performance prédictive des modèles, reprenant le même schéma analytique que celui de l'étude de la qualité d'ajustement, peut être d'utiliser d'autres indicateurs de performance prédictive. L'utilisation de méthodes de validation croisée permettrait également de mieux évaluer la qualité de prédiction (voir paragraphe Qualité de prédiction page 9).

## Exploration des scénarios sylvicoles

La simulation des trois scénarios sylvicoles est répétée 30 fois pour chaque placette appartenant à la série de production de bois. Le temps d'exécution des 30 répétitions, pour les trois scénarios, est d'environ 8 heures par placette, et se fait par calcul parallèle sous R, à l'aide du *package* `parallel` (v4.1.2; Ripley, Tierney et Urbanek, 2011). La densité moyenne, exprimée en nombre de CB par hectare, est construite à partir des 30 répétitions de simulation des trois scénarios sylvicoles pour chaque placette (voir Simulation des scénarios avec Samsara2 page 11). L'enveloppe autour de chaque courbe correspond à l'intervalle de confiance à 95% de cette moyenne estimée.

D'après la FIGURE 9, le comportement des scénarios est assez similaire pour les six placettes. Le scénario de libre évolution du peuplement, sans aucune intervention de coupe, présente une décroissance de la densité moyenne de CB dans le peuplement, notamment pendant la phase transitoire des 200 premières années de la simulation. La diminution de la densité moyenne de CB semble cependant s'atténuer dans les dernières décennies de la simulation. Lorsque l'on compare les deux scénarios impliquant des coupes à celui de libre évolution, on remarque un effet quantitatif et un effet cinétique de l'exploitation. En effet, les scénarios d'exploitation du peuplement présentent une densité moyenne de CB plus faible (effet quantitatif) et une décroissance plus rapide (effet cinétique) que le scénario de libre évolution sur la période. Enfin, la comparaison entre les deux scénarios d'exploitation révèle également des effets quantitatif et cinétique, attribués à l'effort de rétention. La densité moyenne de CB est plus élevée (effet quantitatif) et la décroissance est légèrement moins forte (effet cinétique) pour un effort de rétention s'élevant à 10 arbres par ha que pour une rétention de 3 arbres par ha (scénario BAU).

---

15. (EN) Receiver Operating Characteristic

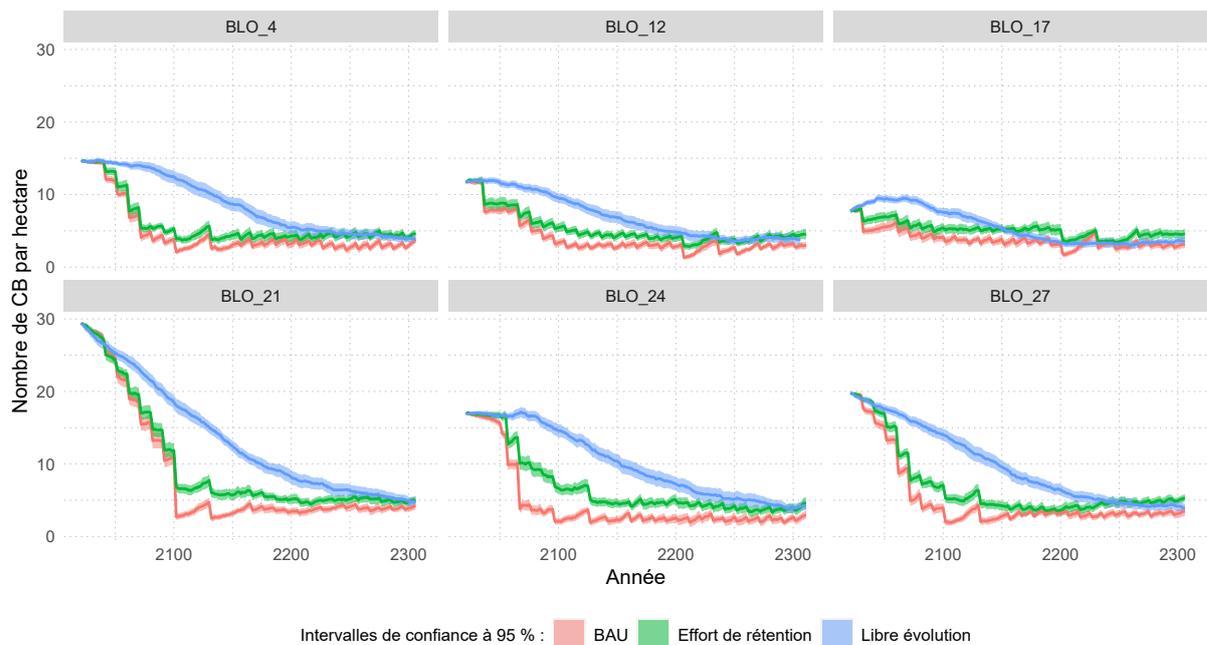


FIGURE 9 – Densité moyenne de CB par placette selon le scénario

## Discussion

Le premier volet se base sur l’hypothèse biologique selon laquelle il existe un processus temporel d’apparition de cavités basses. En supposant une relation croissante âge-diamètre fixe pour une essence dans une placette, on s’attend à une augmentation de la probabilité de présence d’une CB avec le diamètre de l’arbre. La représentation de la probabilité d’apparition de CB selon le type de *priors* (voir FIGURE 7 page 15), semble montrer que le modèle calibré avec des *priors* non informatifs, c.-à-d. sans information extérieure, affiche une fonction de répartition quasi-constante sur la plage des diamètres observés. Ce résultat s’interprète, au niveau biologique, comme le fait qu’il n’existerait pas de processus temporel du développement de CB. Autrement dit, le massif serait composé à 7 % de chênes porteurs de CB, sans processus de disparition ou d’apparition de CB au cours du temps, ce qui est contraire au processus temporel des dendromicrohabitats décrit dans la littérature. À l’inverse, l’apport d’information issue de la base de données européenne dans le modèle permet de décrire un processus temporel d’augmentation de la probabilité d’apparition de CB avec le diamètre, et donc, par analogie, dans le temps. Le résultat porté par ce second modèle est cohérent avec l’hypothèse biologique sur la dynamique d’apparition de cavités basses, tout en étant bien ajusté sur les données de la Grésigne. L’approche bayésienne mobilisant l’information extérieure issue des forêts d’Europe et d’Iran permet de réduire les plages de valeurs prises par les paramètres estimés et de diminuer la corrélation entre eux, au profit d’une meilleure interprétation marginale des paramètres. L’apport d’information européenne facilite l’interprétation biologique du modèle, tout en s’ajustant au cas particulier de la Grésigne.

L’interprétation du paramètre de forme  $k < 1$  ne permet pas de confirmer l’attendu selon lequel le processus de sénescence de l’arbre est favorable à l’augmentation de la probabilité d’apparition de CB dans le contexte de la Grésigne. Ce résultat a des implications importantes en termes de choix d’orientation sylvicole car cela signifie que le recrutement de nouvelles CB dans le peuplement se fait en priorité via des arbres de diamètres faibles à moyens. Néanmoins, ce résultat pourrait potentiellement être en partie artificiel, dû à un phénomène de censure à droite non pris en compte dans la modélisation statistique.

La censure à droite survient si la présence d'une CB affecte la mortalité de l'arbre. Un arbre porteur de CB, du fait de son état de décomposition ligneuse au niveau du tronc et remontant petit à petit dans la tige, pourrait présenter un risque de mortalité naturelle accru (bien que cela n'ait jamais été démontré), ou être préférentiellement éliminé par les gestionnaires du fait de sa perte de valeur marchande.

Comme évoqué plus haut, la qualité d'ajustement reste globalement bonne, quel que soit le type de *priors*. En revanche, les modèles avec les deux types de *priors* montrent une qualité de prédiction plus faible, et une inaptitude à reproduire plusieurs statistiques observées. Le fait que la qualité de prédiction du modèle soit moins bonne que la qualité d'ajustement n'a, par construction, rien d'alarmant. Cette dégradation du potentiel prédictif est ici accentuée par le fait que l'effet aléatoire placette est important et n'est pas ajusté dans la prédiction alors qu'il l'est dans la calibration. De plus le découpage des données en deux sous-échantillons (jeu de calibration et jeu de validation) est particulièrement défavorable à la prédiction. Ce choix englobe l'effet observateur et l'effet de l'année d'acquisition. Cette approche peut donc être qualifiée de conservatrice car elle sous-estime potentiellement la performance prédictive. On aurait pu appliquer une autre approche, comme la validation croisée, sur ces données en effectuant des tirages répétés des placettes de validation dans le corpus 2021-2022. Cette méthode aurait probablement donné une qualité de prédiction plus forte, mais les tendances générales discutées ici seraient probablement restées inchangées.

L'apport d'information issue des autres forêts d'Europe semble renforcer une tendance à sous-estimer le nombre de cavités basses en Grésigne, par rapport au modèle uniquement basé sur les données de la Grésigne (placettes BLO\_22 et BLO\_23 en FIGURE 8 page 15). Le fait que le modèle informé sous-estime toujours plus que le modèle sans information extérieure indique un caractère atypique du massif de la Grésigne par rapport au contexte européen, ce en dépit du fait qu'on a bien intégré la variabilité entre sites à l'échelle de l'Europe dans le *prior* informatif. Il serait donc intéressant de vérifier la représentativité de la Grésigne dans les sites européens afin de savoir si le site de la Grésigne est vraiment singulier ou représenté par les sites d'autres forêts.

Plusieurs pistes sont envisageables pour tenter d'expliquer cette singularité. Il est possible que le régime sylvicole transitoire de la Grésigne vers de la futaie régulière, et anciennement gérée en TSF, affecte la disponibilité des cavités basses dans le temps. D'après une telle hypothèse, l'usage de la référence européenne par l'approche bayésienne permettrait d'anticiper la trajectoire « à la baisse » de la dynamique d'apparition de CB dans un contexte transitoire. Cette anticipation est cruciale pour définir des objectifs de gestion. À l'inverse de cette hypothèse de singularité transitoire, il se peut simplement que cette singularité soit pérenne, auquel cas la conversion de régime sylvicole n'affecterait pas la dynamique d'apparition des cavités basses dans le massif de la Grésigne. En considérant les paramètres estimés avec le modèle informé dans les simulations du second volet du travail, on a pris le parti de considérer cette singularité de la Grésigne comme transitoire.

L'exploration de scénarios de gestion sylvicole par simulation révèle une hiérarchie des effets sur la disponibilité des cavités basses au sein d'un peuplement. Les effets quantitatifs et cinétiques liés à l'exploitation d'un peuplement par rapport à la libre évolution sont plus forts que ceux liés à la variation du nombre d'arbres de rétention. Ce résultat se voit sur les scénarios d'exploitation impliquant des interventions de coupes lorsqu'ils sont comparés avec le scénario de libre évolution, sans intervention. Néanmoins, l'effort de rétention, fixé à 10 arbres par hectare dans la simulation, permet de préserver un plus grand nombre de cavités basses. Malgré une diminution globale de la densité de CB dans les six placettes quel que soit le scénario, l'effort de rétention d'arbres porteurs de

CB ou d'arbres à haut potentiel freine significativement cette diminution, surtout sur la période transitoire jusqu'aux années 2200. Cet effort de préservation par la rétention ne semble cependant pas suffisant, à la vue de la décroissance des scénarios d'exploitation du peuplement par rapport à celle de la libre évolution. Il semble donc nécessaire d'interroger la construction de cet effort de rétention mais également de réfléchir à d'autres méthodes de préservation de la biodiversité associée aux cavités basses, afin d'améliorer l'efficacité de la gestion intégrative. La simulation d'un cycle complet de scénarios de gestion montre une diminution systématique de la disponibilité de CB dans le temps, même pour le scénario de libre évolution. Ce résultat peut potentiellement être imputable à la transition de l'ancien régime de la Grésigne de TSF vers un régime de futaie régulière, similaire aux autres forêts européennes. Cette décroissance de la densité de CB semble cependant s'atténuer vers la fin de la simulation. Une piste d'amélioration serait peut être d'allonger la durée de la simulation afin d'étudier le comportement de la densité moyenne de CB à plus long terme. Il serait également intéressant d'intégrer, dans le modèle de croissance d'arbres de Samsara2, le modèle calibré sur la Grésigne sans information extérieure, afin de voir si cette tendance de baisse de la disponibilité des CB dans le temps est également visible dans la simulation.

## Références

- Christophe BOUGET, Laurent LARRIEU et Antoine BRIN : Key features for saproxylic beetle diversity derived from rapid habitat assessment in temperate forests. *Ecological Indicators*, 36:656–664, janvier 2014. ISSN 1470160X. URL <https://linkinghub.elsevier.com/retrieve/pii/S1470160X13003634>.
- Stephen P. BROOKS et Andrew GELMAN : General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455, décembre 1998. ISSN 1061-8600, 1537-2715. URL <http://www.tandfonline.com/doi/abs/10.1080/10618600.1998.10474787>.
- Benoit COURBAUD, Valentine LAFOND, Guillaume LAGARRIGUES, Ghislain VIEILLEDENT, Thomas CORDONNIER, Franck JABOT et François de COLIGNY : Applying ecological model evaluation : Lessons learned with the forest dynamics model Samsara2. *Ecological Modelling*, 314:1–14, octobre 2015. ISSN 03043800. URL <https://linkinghub.elsevier.com/retrieve/pii/S0304380015002884>.
- Benoit COURBAUD, Laurent LARRIEU, Daniel KOZAK, Daniel KRAUS, Thibault LACHAT, Sylvie LADET, Jörg MÜLLER, Yoan PAILLET, Khosro SAGHEB-TALEBI, Andreas SCHUCK, Jonas STILLHARD, Miroslav SVOBODA et Sergey ZUDIN : Factors influencing the rate of formation of tree-related microhabitats and implications for biodiversity conservation and forest management. *Journal of Applied Ecology*, 59(2):492–503, février 2022. ISSN 0021-8901, 1365-2664. URL <https://onlinelibrary.wiley.com/doi/10.1111/1365-2664.14068>.
- Benoit COURBAUD, Cyrille PUPIN, Anthony LETORT, Alain CABANETTES et Laurent LARRIEU : Modelling the probability of microhabitat formation on trees using cross-sectional data. *Methods in Ecology and Evolution*, 8(10):1347–1359, octobre 2017. ISSN 2041-210X, 2041-210X. URL <https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12773>.
- Matthew J. DENWOOD : **runjags** : An R Package Providing Interface Utilities, Model Templates, Parallel Computing Methods and Additional Distributions for MCMC Models in **JAGS**. *Journal of Statistical Software*, 71(9), 2016. ISSN 1548-7660. URL <http://www.jstatsoft.org/v71/i09/>.
- Jean-Jacques DUCASSE et Hervé BRUSTEL : « Ah, cette grésigne ! » : de Rabil à Natura 2000. *Revue Forestière Française*, Sup10(3), 2008. ISSN 1951-6827, 0035-2829. URL <https://hal.archives-ouvertes.fr/hal-03449668>.
- Samuel DUFOUR-KOWALSKI, Benoît COURBAUD, Philippe DREYFUS, Céline MEREDIEU et François de COLIGNY : Capsis : an open software framework and community for forest growth modelling. *Annals of Forest Science*, 69(2):221–233, mars 2012. ISSN 1286-4560, 1297-966X. URL <https://annforsci.biomedcentral.com/articles/10.1007/s13595-011-0140-9>.
- Daniel KRAUS et Andreas SCHUCK : Wise use of our forests : the integrative approach, 2017. URL <https://integrateplus.eu/>.
- Laurent LARRIEU : Les dendromicrohabitats (dmhs) : des éléments clés pour la biodiversité. In *Formation aux dendromicrohabitats des services verts de la ville de Pau*, page 39 p., Pau, France, mars 2019. URL <https://hal.inrae.fr/hal-02786337>.

OFFICE NATIONAL DES FORÊTS : Document d'objectif du site « Forêt de la Grésigne » FR7300951 – Synthèse, décembre 2005. URL <https://www.occitanie.developpement-durable.gouv.fr/docob-foret-de-la-gresigne-a24773.html>.

Jean RABIL : Ah, cette Grésigne! Catalogue des coléoptères de la forêt de la Grésigne (Tarn). *Nouvelles archives du Muséum d'histoire naturelle de Lyon*, 29(1):3–174, 1991. ISSN 0373-6636. URL [https://www.persee.fr/doc/mhnly\\_0373-6636\\_1991\\_num\\_29\\_1\\_1100](https://www.persee.fr/doc/mhnly_0373-6636_1991_num_29_1_1100).

Aurélien VUIDOT, Yoan PAILLET, Frédéric ARCHAUX et Frédéric GOSSELIN : Influence of tree characteristics and forest management on tree microhabitats. *Biological Conservation*, 144(1):441–450, janvier 2011. ISSN 00063207. URL <https://linkinghub.elsevier.com/retrieve/pii/S0006320710004295>.

	Diplôme : Master Spécialité : Mathématiques appliquées, statistique Spécialisation / option : Science des données pour la biologie Enseignant référent : François HUSSON
Auteur(s) : Pierre COTTAIS Date de naissance* : 10/07/1998	Organisme d'accueil : Centre INRAE Occitanie-Toulouse (UMR Dynafor)
Nb pages : 21      Annexe(s) :	Adresse : 24 Chemin de Borde-Rouge, 31326 Castanet-Tolosan
Année de soutenance : 2022	Maître de stage : Fabien LAROCHE
Titre français : Effets de mesures de conservation intégrative en forêt sur la dynamique d'apparition d'un dendromicrohabitat	
Titre anglais : Effects of integrative conservation measures on the dynamics of a tree-related microhabitat formation	
Résumé (1600 caractères maximum) : La cavité basse (CB) de tronc est un dendromicrohabitat dans lequel vit un grand nombre d'organismes saproxyliques indispensables au cycle de vie du bois, notamment des coléoptères comme le Taupin violacé, une espèce menacée. L'objectif de ce travail est d'explorer des scénarios de gestion sylvicole intégrative par simulation pour en mesurer les effets sur la disponibilité des CB en forêt de Grésigne (Tarn), connue pour sa forte richesse entomologique. Cette simulation intègre la dynamique d'apparition de CB, modélisée par une approche bayésienne permettant de mobiliser des connaissances extérieures issues d'une base de données de forêts d'Europe et d'Iran. Il résulte de cet apport d'information <i>a priori</i> , une sous-estimation du nombre de CB par le modèle sur les données de validation. Cela est potentiellement imputable à la dynamique de transition de la Grésigne, en matière de régime sylvicole, illustrant l'aspect singulier de ce massif, par rapport au contexte européen, en matière de disponibilité d'habitats pour les organismes saproxyliques. La simulation révèle une tendance à la baisse de la densité de CB, quelle que soit la gestion pratiquée. L'effort de rétention d'arbres porteurs de CB ou à haut potentiel (gros arbres) semble néanmoins freiner cette baisse de la disponibilité de CB dans le temps. Le fait que la densité de CB diminue, même pour un scénario de libre évolution du peuplement, illustre les implications quantitatives de transition de la Grésigne vers un régime similaire aux autres forêts européennes.	
Abstract (1600 caractères maximum) : The trunk cavity is a tree-related microhabitat for a lot of saproxylic organisms, especially beetles such as the <i>Limoniscus violaceus</i> , a threatened species essential to the life cycle of wood. We explore silviculture management scenarios using stand simulation in order to quantify the effects on trunk cavities in the Grésigne forest (Tarn, France), well known for its richness and diversity of insects. The simulation is based on a cavity formation model, estimated with a Bayesian approach using information from European forests database. The use of informative priors in the model leads to underestimating the number of cavities in Grésigne forest. It might be caused by the ongoing transition of the forestry management, illustrating the fact that the forest is entomologically distinctive by its number of cavities, compared to the other European managed forests. The simulation shows negative effects of forestry management on the number of cavities. However, the retention of cavity bearing trees and big trees without cavity reduces this decrease in the number of cavities. The decrease of cavities even for the unmanaged stand illustrate the putative quantitative implications in terms of cavities density of Grésigne transition towards a forest management regime used in other European forests.	
Mots-clés : analyse de survie, cavité basse, inférence bayésienne, modèle individu-centré, chaînes de Markov, rétention d'arbres	
Key Words: Bayesian inference, individual based model, Markov chain, retention forestry, survival analysis, trunk cavity	

\* Élément qui permet d'enregistrer les notices auteurs dans le catalogue des bibliothèques universitaires