



Rapport de stage
Mathématiques Appliquées Statistique, M2 MEA

Effectué à l'**INRAE**

Analyse de viabilité d'un modèle mécaniste pour l'aide à la décision de gestion : exemple d'un peuplement de chêne

Michelle MALARA

Tuteurs :

Marion JOURDAN (UMR SILVA)
Jean-Baptiste PICHANCOURT (UR LISC)
Jean-Denis MATHIAS (UR LISC)

Remerciements

À l'issue de ce stage, je tiens à remercier mes tuteurs pour leur confiance, leur bienveillance et tout ce qu'ils m'ont apporté au cours de ce stage. Merci Jean-Denis pour ton enthousiasme et tes encouragements. Merci Jean-Baptiste pour ta bonne humeur quotidienne et tes multiples anecdotes insolites. Merci Marion pour ta convivialité et ta patience lors de nos longues réunions en visio.

Je remercie également toute l'équipe du LISC pour l'ambiance accueillante et chaleureuse lors des poses-café ainsi que pour tous les échanges enrichissants qui nous a été donné d'avoir. Enfin, j'adresse mes remerciements aux autres stagiaires pour tous les moments que nous avons partagés.

Table des matières

| | |
|--|-----------|
| Introduction | 1 |
| 1 Modèle de trouées forestières : ForCEEPS | 3 |
| 1.1 Modèle environnement | 3 |
| 1.1.1 Données géographiques | 3 |
| 1.1.2 Variables thermiques | 4 |
| 1.1.3 Bilan hydrique | 4 |
| 1.2 Modèle des plantes | 4 |
| 1.2.1 Modèle de la forme du houppier | 5 |
| 1.2.2 Modèle de démographie à l'échelle individuelle | 5 |
| 1.2.3 Variables à l'échelle populationnelle | 7 |
| 2 Problème de gestion sylvicole étudié | 8 |
| 2.1 Système : la forêt | 8 |
| 2.2 Contrôle : la gestion | 9 |
| 2.2.1 Variables de contrôle | 9 |
| 2.2.2 Interventions | 9 |
| 2.3 Contraintes : les objectifs écologiques et économiques | 11 |
| 3 Méthode de résolution du problème de gestion adapté aux contraintes | 12 |
| 3.1 Résolution par algorithme génétique | 13 |
| 3.1.1 Description générale de l'algorithme génétique | 13 |
| 3.1.2 Définition des individus de l'algorithme génétique | 15 |
| 3.1.3 Opérations réalisées dans l'algorithme génétique | 15 |
| 3.1.4 Fonction objectif | 16 |
| 3.1.5 Analyse de sensibilité de l'algorithme | 18 |
| 3.2 Classification des résultats | 18 |
| 3.2.1 Analyse en composantes principales | 18 |
| 3.2.2 Classification à ascendance hiérarchique | 20 |
| 4 Mise en application | 20 |
| 4.1 Cas d'études | 20 |
| 4.2 Présentation des résultats de l'algorithme | 21 |
| 4.2.1 Analyse en composantes principales | 21 |
| 4.2.2 Classification à ascendance hiérarchique | 23 |
| 5 Discussion | 26 |
| 5.1 Analyse des résultats | 26 |
| 5.1.1 Résultats de l'algorithme génétique | 26 |
| 5.1.2 Résultats de l'ACP | 26 |
| 5.1.3 Analyse des itinéraires de gestion | 26 |
| 5.2 Discussion méthodologique | 27 |
| Conclusion | 28 |

| | | |
|----------|---|------------|
| A | Équations de ForCEEPS | I |
| A.1 | Modèle environnement | I |
| A.1.1 | Bilan hydrique | I |
| A.1.2 | Variables climatiques | II |
| A.2 | Modèle des plantes | III |
| A.2.1 | Modèle du houppier | III |
| A.2.2 | Modèle de croissance | IV |
| A.2.3 | Modèle de mortalité | VII |
| A.2.4 | Modèle d'établissement | VIII |
| B | Modèle ForClim | X |
| B.1 | Modèle environnement | X |
| B.2 | Modèle des plantes | X |
| B.3 | Modèle du sol | XI |
| B.4 | Les limites | XI |
| C | Test du nombre de classe pour la CAH | XII |
| D | Figures complémentaires de l'analyse des résultats | XIV |

Introduction

L'INRAE (Institut National de Recherche pour l'Agriculture, l'alimentation et l'Environnement) est issu de la fusion entre l'INRA (Institut National de la Recherche Agronomique) et l'IRSTEA (Institut national de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture) qui a eu lieu en 2020. Cet institut de recherche publique français a pour mission de réaliser des travaux de recherche scientifique et technologique dans les domaines de l'agronomie, la gestion durable des écosystèmes et du territoire et de l'économie circulaire. Ce stage est issu d'un travail commun entre deux des différentes unités de recherche qui compose l'INRAE.

La première, l'équipe EcoSilva (écologie des forêts et des écosystèmes peu anthropisés), est issue de l'unité mixte de Recherche Silva. Sa mission est, entre autres, l'étude de l'adaptation et de la résilience des arbres et des écosystèmes aux perturbations dues aux changements environnementaux, principalement sur des périodes longues. Le second, le LISC (Laboratoire d'Ingénierie pour les Systèmes Complexes), est spécialisé dans l'étude et la modélisation des systèmes dynamiques. Cette unité de recherche a développé des approches quantitatives pour aider la gestion durable de systèmes socio-écologiques.

Le changement climatique rend plus vulnérables les écosystèmes en augmentant les risques d'aléas climatiques. Cela implique automatiquement une augmentation de la vulnérabilité des populations humaines qui dépendent des services fournis par les écosystèmes. Les forêts sont un exemple parlant de cette interdépendance entre les besoins des sociétés humaines et le fonctionnement des écosystèmes.

Maintenir dans le temps la fourniture des services écosystémiques forestiers est donc un enjeu central pour garantir. Il est important de savoir si la gestion actuelle permettra cela et, si ce n'est pas le cas, il faut comprendre comment l'adapter de manière pertinente.

Cependant, il est difficile d'intégrer de manière exhaustive les processus forestiers et les services associés. À cause de cela la prise de décision pour la gestion n'est pas toujours évidente. D'autant plus que les résultats des actions et du changement climatique sur ces écosystèmes ne sont, le plus souvent, mesurables qu'à long terme.

L'idée de ce stage est de développer un outil d'aide à la décision sylvicole afin de maintenir les services de la forêt tout en préservant le fonctionnement de cet écosystème dans un contexte de changement climatique. Pour y répondre, la méthode que nous avons choisie est d'utiliser un modèle mécaniste en le combinant avec des méthodes issues de la théorie du contrôle. L'utilisation d'un problème du contrôle pour étudier le maintien d'un système dans des contraintes en grande dimension n'est pas très répandue dans la littérature. Nous avons exploré lors de ce stage une solution possible.

Il existe un grand nombre de modèles représentant les écosystèmes forestiers. Le modèle que nous avons utilisé, ForCEEPS (Forest Community Ecology and Ecosystem Processes) [1], est un modèle de trouée forestière (ou gap model). Ces modèles permettent notamment l'étude de la succession des différentes espèces végétales. Ce sont des modèles mécanistes, c'est-à-dire qu'ils cherchent à représenter les mécanismes régissant le fonctionnement des écosystèmes. Certains sont à l'échelle du peuplement et traduisent l'état de l'ensemble des arbres présents sur la surface étudiée. D'autres traitent l'échelle individuelle en modélisant chaque arbre indépendamment. Les seconds peuvent aussi étudier des propriétés au niveau de la forêt en utilisant des variables individuelles pour calculer des variables populationnelles.

Généralement, ce sont des modèles où le temps est discrétisé par année.

Dans les modèles de peuplement, ceux étudiant la forêt présentent quelques particularités. D’abord, le mouvement des spécimens n’est pas pris en compte. Ensuite, la notion de croissance des individus a plus d’importance car, en plus du nombre d’arbres, les gestionnaires forestiers s’intéressent à la quantité de bois produits. Le fonctionnement des modèles de peuplement forestier est donc basé sur la simulation de la croissance, la régénération et la mortalité des arbres. Ces mécanismes sont influencés par différentes contraintes, qu’elles soient inhérentes à l’individu ou exogène.

Ces modèles cherchent à représenter fidèlement un système complexe et tous les mécanismes rentrant en jeu dans sa dynamique. En plus de cela, la succession forestière est un phénomène visible sur temps long. Les simulations pouvant dépasser les centaines d’années, il est préférable qu’une itération ne soit pas trop coûteuse. Il est nécessaire de faire des choix de modélisation pour trouver la complexité permettant une certaine précision tout en ayant un temps de calcul acceptable.

Lors de ce stage notre choix se porte sur le modèle ForCEEPS car il est considéré comme un compromis entre les modèles complexes et ceux plus généraux utilisés pour étudier le comportement des forêts. De plus c’est un modèle calibré pour les espèces arboricoles d’Europe de l’Ouest [1]. Pour finir, il est adapté à la modélisation des forêts mixtes en intégrant la gestion et le changement climatique sur temps long [2]. J’analyse dans la section 1 l’ensemble de sa dynamique grâce aux travaux de la littérature [1, 3, 4] et aux codes de la plateforme partagée CAPSIS [5].

En s’appuyant sur les simulations avec le modèle ForCEEPS, nous voulons explorer l’espace des possibles pour identifier des pistes pouvant nourrir la réflexion des gestionnaires. Dans cette optique, nous utilisons la théorie du contrôle [6, 7, 8, 9]. La définition d’un problème de contrôle se fait en trois étapes qui sont détaillées dans la section 2. La dynamique et les variables d’états du système sont identifiées grâce à l’analyse du modèle faite en amont. Nous nous basons sur un cas concret d’application pour définir l’ensemble des contrôles, permettant d’agir sur la dynamique du système, ainsi que l’ensemble des contraintes, caractérisant si l’état de la forêt est satisfaisant ou non.

En considérant notre problème comme un problème d’optimisation en grande dimension, il est possible d’explorer les possibilités pour identifier des politiques d’action satisfaisant les contraintes. Nous utilisons un algorithme génétique pour résoudre ce problème. La description de l’algorithme que j’ai implémenté ainsi que la définition du problème génétique basé sur notre problème de contrôle est faite dans la section 3.

Pour finir, l’utilisation des méthodes d’analyse en composantes principales et de classification à ascendance hiérarchique faites dans la section 4 nous permet de classer les politiques d’action identifiées par l’algorithme génétique et d’identifier différents types d’itinéraire de gestion viable.

1 Modèle de trouées forestières : ForCEEPS

ForCEEPS est un modèle mécaniste individu-centré présenté en 2021 [1]. C'est un modèle calibré pour les principales espèces d'arbres en France. Il a déjà été utilisé pour de la prévision de l'évolution à long terme des forêts mélangées en intégrant la gestion et le changement climatique [2].

Son fonctionnement, schématisé sur la figure 1, est dérivé du principe du modèle ForClim [3] décrit dans l'annexe B. Le détail des équations implémentées dans ForCEEPS est donné dans l'annexe A. Il est divisé en modules qui peuvent être séparés en deux sous-modèle. Le sous-modèle environnement contient les données d'entrée, les variables thermiques et le bilan hydrique et celui des plantes correspond au reste des processus. Les interactions entre les modules se font grâce à des variables. Les calculs se font chaque année pour un sous-ensemble de la forêt appelé placette.

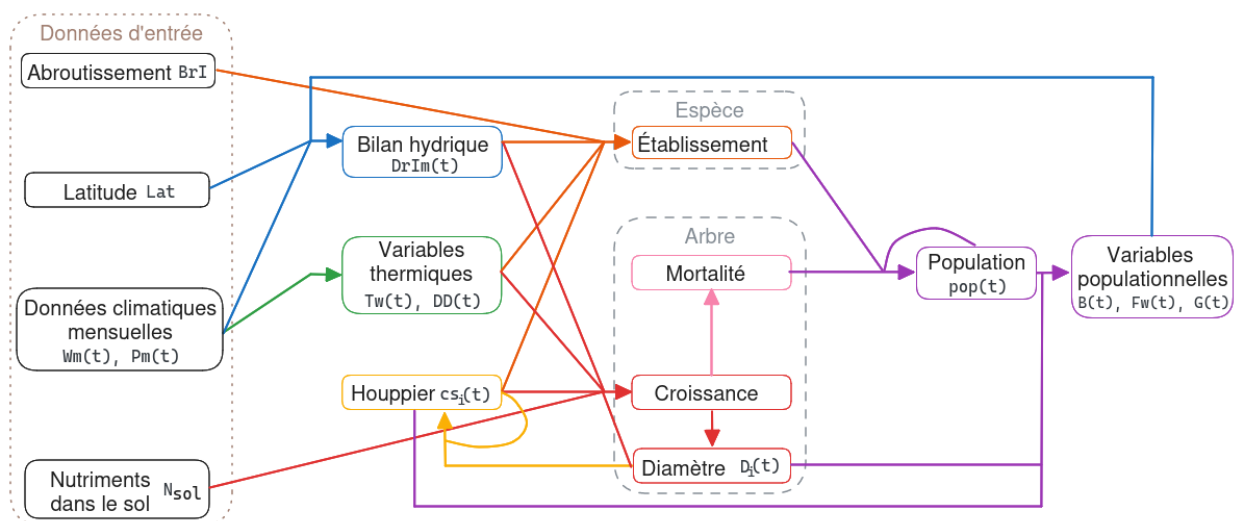


FIGURE 1 – Schéma de l'organisation du modèle ForCEEPS. Les flèches représentent des interactions via des variables, les couleurs représentent les différents processus (vert : calcul des variables thermiques, bleu : bilan hydrique, jaune : modélisation du houppier, rouge et rose : croissance et mortalité des arbres, orange : établissement des espèces, violet : calcul des variables populationnelles), les pointillés gris donnent le niveau de calcul des processus, s'il n'est pas précisé, c'est un processus de niveau global et les points beiges représentent les données d'entrée du modèle.

1.1 Modèle environnement

C'est dans le modèle environnement que sont représentées les contraintes extérieures influençant l'évolution des plantes. Ici, je les ai regroupées en trois catégories. Les données géographiques sont des données d'entrée non transformées caractérisant le site étudié. Les variables thermiques sont calculées à partir des données climatiques utilisées lors de la simulation. Les contraintes liées à l'eau sont calculées dans le module bilan hydrique.

1.1.1 Données géographiques

Il existe plusieurs contraintes liées au site représentées par des données d'entrée. Elles sont constatées dans le temps et identiques en tout point de la placette simulée. La première

est la quantité d’azote disponible N_{soil} . Elle caractérise la fertilité du sol. Il y a aussi la latitude Lat qui permet de calculer la durée des jours durant l’année ce qui influence le bilan hydrique. La dernière contraintes est dû à la présence des animaux. En effet, ils se nourrissent de jeunes pousses de certaines espèces et ont donc nécessairement une influence sur leur développement. La variable d’intensité d’abrutissement BrI est utilisée pour caractériser l’influence des animaux.

1.1.2 Variables thermiques

Les variables thermiques utilisées dans le modèle sont définies à l’échelle annuelle, mais leur calcul est fait avec des données brutes issues des relevés climatiques définis à l’échelle mensuelle (voir en vert sur la figure 1). Il est possible d’utiliser dans ForCEEPS des données historiques ou des scénarios de prédiction comme les RCPs (Representative Concentration Pathway) du GIEC (Groupe d’experts intergouvernemental sur l’évolution du climat ou IPCC Intergovernmental Panel on Climate Change) [10].

Pour chaque année $t = 1, 2, \dots, T$, les moyennes des températures $T_m(t)$ sachant le mois $m = 1, \dots, 12$ sont utilisées pour calculer [3, 11] :

- la température hivernale $T_W(t)$ comme la moyenne des températures des mois de décembre, janvier et février.
- le degré-jour $DD(t)$ comme la somme des températures de chaque mois.

Ces variables permettent d’évaluer l’impacte des températures sur la croissance et l’établissement des arbres.

1.1.3 Bilan hydrique

Le dernier module environnemental permet de calculer l’eau disponible pour les arbres à partir des données mensuelles de précipitations $W_m(t)$ et de températures $T_m(t)$ et de la durée journalière évaluer grâce à la donnée géographique Lat (voir en bleu sur la figure 1). Ces données ainsi que la quantité de végétation calculée dans le modèle des plantes (voir section 1.2.3) permettent d’évaluer l’évapotranspiration des arbres $Et_m(t)$ et du sol $D_m(t)$. C’est ainsi que sont obtenu les indices de stress hydrique mensuel $DrI_m(t)$ puis annuel $DrI(t)$ [1]. La définition de ces variables permet en suite de quantifier les perturbations que subiront les arbres à cause du manque d’eau.

1.2 Modèle des plantes

ForCEEPS est un modèle individu-centré de dynamique de peuplement. Par conséquent, les mécanismes décrits ici dépendent de chaque arbre modélisé. Une des particularités de ForCEEPS, comparé aux autres modèles de trouées forestières, est sa représentation du feuillage. l’attention particulière portée à la forme du houppier permet une modélisation plus réaliste des mécanismes de compétition pour la lumière.

L’état de la forêt est modélisé année après année. L’intégralité des variables à l’échelle d’une population $pop(t)$ peut être calculée à partir de l’ensemble des variables définies à l’échelle individuelle (diamètres $D_i(t)$ et proportion du houppier $cs_i(t)$).

1.2.1 Modèle de la forme du houppier

Le feuillage a un rôle important dans la vie des arbres notamment car c'est lui qui permet la photosynthèse. Il faut aussi prendre en compte celui des arbres qui l'entoure car leur feuillage bloque une partie de la luminosité disponible. La représentation du feuillage à travers la forme de son houppier aura une influence sur les résultats du modèle (voir en jaune sur la figure 1).

ForCEEPS est individu-centré ce qui permet de prendre en compte indépendamment le houppier de chaque arbre. Sa représentation géométrique est plus réaliste que dans d'autres modèles : c'est un cône tourné vers le bas, voir figure 2. La coupe verticale du cône forme un triangle d'une aire de 1 m^2 . L'espace vertical est discrétisé en couche, ou strate, de $\Delta x = 1 \text{ m}$, voir figure 2. On peut estimer la surface foliaire $LA_{i,x}(t)$ de chaque arbre i dans chaque strate x à partir de sa surface foliaire totale $LA_i(t)$ qui est elle-même calculée à partir de $D_i(t)$ et de $cs_i(t)$. On définit aussi des coefficients $r_{i,x}(t)$ comme la surface des parallélépipèdes formés par l'intersection de ce triangle et des couches.

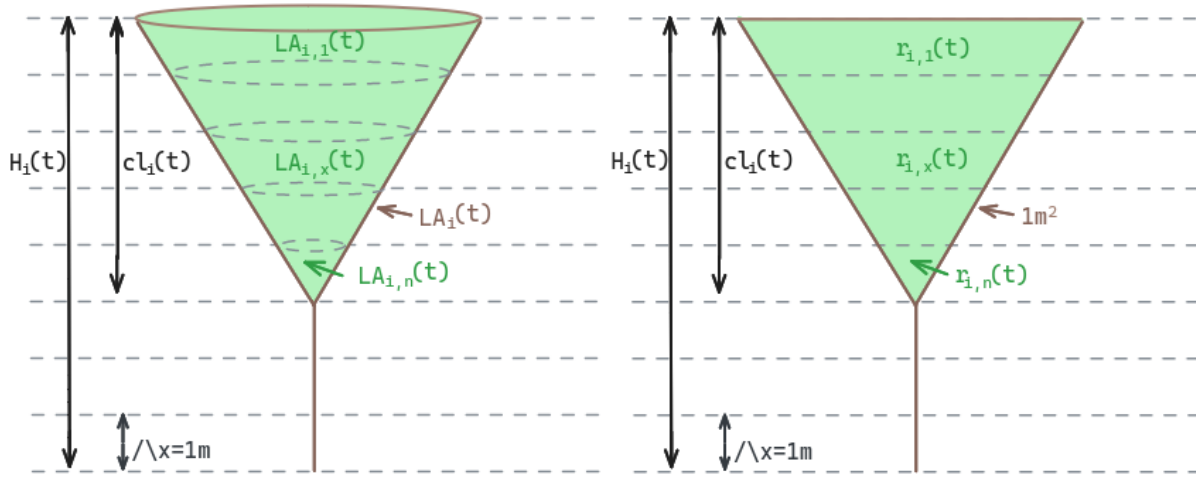


FIGURE 2 – Schéma de la forme du houppier d'un arbre i dans ForCEEPS (à gauche) et coupe verticale pour le calcul des coefficients de strate $r_{i,x}(t)$ (à droite) sachant sa taille $H_i(t)$, celle de son houppier $cl_i(t)$, sa surface foliaire $LA_i(t)$, la surface foliaire $LA_{i,x}(t)$ appartenant à la strate x .

La taille du houppier $cl_i(t)$ de chaque arbre et plus précisément le rapport entre la taille de l'arbre et celui de son houppier $cs_i(t) = cl_i(t)/H_i(t)$ est influencé par le feuillage des arbres qui l'entourent [12]. Ces variables permettent d'estimer la lumière accessible $Lav_x(t)$ dans chaque strate. Cela permet de calculer la quantité de lumière disponible pour chaque arbre et donc leur croissance et indirectement leur mortalité. L'établissement de nouveaux individus est aussi impacté par la lumière disponible au sol. Ces mécanismes sont détaillés dans la section 1.2.2.

1.2.2 Modèle de démographie à l'échelle individuelle

Ce module regroupe les trois mécanismes des peuplements forestiers cités plus tôt : l'établissement, la croissance et la mortalité. L'ensemble de ces mécanismes sont basés sur des données d'espèces paramétrées grâce aux données de la littérature [1]. La majorité des espèces arboricoles françaises et de l'Ouest Européens sont paramétrées dans ForCEEPS.

1.2.2.1 Modèle de croissance

La croissance d'un arbre (voir en rouge sur la figure 1) est déterministe. Elle dépend des variables climatiques et géographiques et de son espèce. On s'intéresse à la croissance du diamètre à hauteur de poitrine $D(t)$. Son calcul est fait en deux étapes. D'abord, la croissance optimal $\Delta D_{\text{opt}_i}(t)$ est calculée. Elle varie positivement avec le diamètre. La loi qu'elle suit dépend de l'espèce de l'arbre.

Il faut ensuite modifier le taux optimal en fonction des contraintes comme la sécheresse, les températures et la quantité de nutriment disponible représentées respectivement par les variables $DrI(t)$, $DD(t)$ et N_{soil} calculées dans le modèle environnement (voir section 1.1). L'influence de la température sur l'assimilation du CO_2 est aussi étudiée. Pour finir, le feuillage a deux effets sur la croissance. Le premier est dû à l'ombre créée par les autres arbres. L'effet des feuilles au-dessus de l'arbre est pondéré par les coefficients $r_{i,x}(t)$ définis plus tôt (voir section 1.2.1). Deuxièmement, la lumière captée par un arbre influence sa croissance. Elle est directement liée à la forme de son feuillage et plus précisément à la proportion du houppier.

La croissance réelle du diamètre $\Delta D(t)$ durant l'année t est ainsi obtenue et permet le calcul du nouveau diamètre.

$$D(t+1) = D(t) + \Delta D(t)$$

1.2.2.2 Modèle de mortalité

Contrairement à la croissance, la mortalité (voir en rose sur la figure 1) n'est pas déterministe mais partiellement stochastique. Un taux de mortalité est déterminé pour chaque arbre. Il est ensuite comparé à la résultante d'une variable aléatoire. L'ensemble des individus morts l'année t est noté $M(t)$. Ce taux est la résultante de deux effets.

Le premier est la mort naturelle de l'arbre dû à son vieillissement. Elle est simplifiée par un taux constant qui dépend de l'âge maximum de l'espèce. La seconde est relative à la pression que subit un arbre à cause de la concurrence ou des contraintes extérieures. Elle suit une loi inversement proportionnelle à sa croissance. Pour la calculer la croissance de l'arbre est gardée en mémoire et si elle est considérée trop faible le taux de mortalité augmente. En effet, quand un arbre grandit trop lentement, c'est un signe que les conditions ne lui sont pas favorables.

1.2.2.3 Modèle d'établissement

ForCEEPS ne modélise pas directement le processus de reproduction. La raison est empirique et est liée particulièrement au fait qu'il est difficile de suivre l'ensemble des graines produites et de calculer leur probabilité de survie et de germination dans le sol. Par conséquent, seuls les jeunes arbres de plus de $D_i(t) = 1.27$ cm de diamètre sont comptabilisés dans la modélisation. À partir de cela, ForCEEPS modélise le taux d'établissement comme un seuil, fonctionnant comme des filtres environnementaux (voir en orange sur la figure 1). Par exemple, chaque espèce a un seuil de sécheresse et si la sécheresse de l'année $DrI(t)$ le dépasse alors elle ne pourra pas se régénérer cette année-là. $T_W(t)$ ou $DD(t)$ sont aussi utilisés pour exclure les espèces qui ne tolèrent pas les températures trop froides. De même, le feuillage des arbres diminue la lumière au sol et empêche l'établissement des espèces exigeantes en lumière. Enfin, les animaux se nourrissent de jeunes pousses de certaines espèces leur présence peut donc rendre impossible la régénération. Cette influence est caractérisée par une intensité d'abrouissement BrI constante sur tout le site et dans le temps.

Si les conditions sont favorables à une espèce, cela ne garantit pas pour autant son établissement. En effet, la germination et la survie de la jeune plante peuvent être compromises par d'autres facteurs qui ne sont pas modélisés explicitement. La probabilité que l'espèce s'établisse en conditions favorables est fixe et définie à 10% à partir de données empiriques [1]. Le nombre d'individus établis $nb_{Es}(t)$ appartenant à l'espèce s est tiré selon une loi aléatoire uniforme qui dépend aussi de paramètres spécifiques. Ensuite, chaque individu établi intègre la simulation avec un diamètre initial $D_i(t) = 1.27$ cm. À ce niveau, il est possible de définir la variable $E(t)$ qui représente l'ensemble des individus, toutes espèces confondues, qui se sont établis l'année t .

Ce calcul de base peut être supplémenté par d'autres options. Par exemple, il est possible de prendre en compte la composition de la forêt de l'année précédente pour définir la dispersion des graines en fonction des espèces présentes et de leur proportion. Dans le cadre de ce stage, nous nous intéresserons à l'établissement de base sans option supplémentaire.

1.2.3 Variables à l'échelle populationnelle

Lors des itérations, le modèle calcule des processus à l'échelle individuelle. Les caractéristiques de la forêt sont évaluées pour chaque année t de la simulation, grâce aux propriétés de tous les arbres de la placette modélisée (voir en violet sur la figure 1).

D'abord, l'ensemble des individus présents l'année $t + 1$ est calculé en ajoutant les nouveaux arbres établis $E(t)$ et en enlevant ceux qui sont morts $M(t)$ au courant de l'année.

$$\text{pop}(t + 1) = E(t) \cup \text{pop}(t) \setminus M(t) \quad (1a)$$

Ensuite, deux variables de la biomasse permettent d'évaluer la potentielle production de la forêt. Leurs calculs sont les suivants : $B(t)$ est constituée du bois de la forêt (tronc, branches, brindilles, etc...). La quantité de bois chaque arbre $B_i(t)$ est estimée grâce à son diamètre $D_i(t)$.

$$B(t) = \sum_{\forall i \in \text{pop}(t)} B_i(t) \text{ avec } B_i(t) = 0.12 \times D_i(t)^{2.4} \quad (1b)$$

La deuxième variable $Fw(t)$ représente le poids du feuillage total. Le calcul du poids du feuillage d'un arbre est fait grâce à son diamètre $D_i(t)$ et la proportion de son houppier $cs_i(t)$. Il dépend aussi de son espèce à travers les paramètres a_s , f'_s et f''_s qui caractérisent la forme de son houppier.

$$Fw(t) = \sum_{\forall i \in \text{pop}(t)} Fw_i(t) \text{ avec } Fw_i(t) = f''_s \times cs_i(t) \times D_i(t)^{a_s} = \frac{f''_s}{f'_s} \times LA_i \quad (1c)$$

Cette quantité qui est utilisée dans du calcul du bilan hydrique (voir section 1.1.3).

Enfin, la surface terrière $G(t)$ est une métrique souvent utilisée. Pour un arbre c'est la surface de la coupe du tronc à hauteur de poitrine, elle est donc calculée grâce au diamètre $D_i(t)$. Elle est normalisée par la surface utile Su .

$$G(t) = \sum_{\forall i \in \text{pop}(t)} G_i(t) \text{ avec } G_i(t) = \frac{D_i(t)^2}{Su} \times \frac{\pi}{4} \quad (1d)$$

L'amélioration que propose ForCEEPS, par rapport à ForClim, le modèle dont il est inspiré, est sa représentation plus réaliste des houppiers. L'accessibilité à la lumière influence la croissance, la mort et l'établissement des plantes. Une meilleure simulation individuelle de la compétition lumineuse permet donc de meilleurs résultats notamment pour les forêts mélangées. Cependant, cela entraîne nécessairement une augmentation des calculs ce qui est en partie compensée par la suppression du modèle du sol par rapport à ForClim.

Dans la suite, je décrirais le problème de contrôle associé à ce modèle. Puis je parlerais de la méthodologie adoptée pour sa résolution. Enfin, j'exposerais les résultats obtenus avec l'application étudiée lors du stage.

2 Problème de gestion sylvicole étudié

Le problème de contrôle étudié doit tenir compte du contexte de changement climatique, de l'impact des actions humaines sur l'évolution de l'état de la forêt sur le long terme. La définition d'un problème de contrôle se fait en plusieurs étapes [7, 8, 9] :

1. Il faut commencer par identifier les variables d'état de notre système et leur ensemble de définition X ainsi que la dynamique f qui les régit : $x(t) \in X \quad f : X \longrightarrow X$.
2. Ensuite, il faut caractériser l'ensemble des contrôles U qui nous permettent d'agir sur le système en influençant sa dynamique : $u(t) \in U \quad f : X \times U \longrightarrow X$.
3. Pour finir, un ensemble des contraintes $K \subset X$ doit être déterminé pour définir si un état $x(t) \in K$ est considéré comme satisfaisant ou non.

Les deux derniers dépendent des moyens d'action des acteurs et de leurs objectifs. Nous nous sommes basés sur un cas concret pour les définir.

Le problème de gestion que nous avons étudié est le suivant : en partant d'un état initial fixe $x_0 \in K$, nous souhaitons identifier l'ensemble des suites de contrôle $u(.) \in U$ qui permettent de toujours respecter les contraintes K . Cet ensemble s'appellera dans la suite $V_K(x_0)$ l'ensemble des contrôles satisfaisants au départ notre état x_0 .

$$V_K(x_0) = \{u(.) \in U \mid x(0) = x_0, \forall t \ x(t+1) = f(x(t), u(t)) \in K\} \quad (2)$$

2.1 Système : la forêt

Le système d'intérêt est la forêt comme elle est opérationnalisée classiquement par les gestionnaires forestiers. L'étude des équations, décrites dans la section 1 et détaillée dans l'annexe A, permet d'identifier l'ensemble des variables d'état $x(t)$ de notre système. Il s'agit de l'ensemble des diamètres $D_i(t)$ et des proportions du houppier $cs_i(t)$ de tous les arbres de la population $\text{pop}(t)$.

$$x(t) = \{D_i(t), cs_i(t), \forall i \in \text{pop}(t)\} \quad (3)$$

Cela signifie que toutes les autres variables utilisées peuvent être définies seulement avec $x(t)$. De plus, l'évolution des variables d'état ne dépend que d'elles-mêmes et de la dynamique f du système définie par l'ensemble des équations (voir annexe A).

$$x(t+1) = f(x(t)), \forall t = 1, \dots, T \quad (3a)$$

2.2 Contrôle : la gestion

Il est possible d'appliquer un contrôle $u(t)$ sur le système qui va influencer sa dynamique f et donc son évolution au cours du temps. Ainsi (3a) devient :

$$x(t+1) = f(x(t), u(t)) \quad (3b)$$

Dans le cadre de la gestion forestière, le contrôle appliqué au système peut se résumer par des interventions de coupe et de reboisement. Ici, nous nous concentrerons seulement sur la coupe. Il faut donc définir l'ensemble des arbres $C(t)$ coupés lors de ces interventions. Ces dernières sont caractérisées par des variables de contrôle. Pour les définir, je me suis basée sur l'implémentation des interventions sylvicoles déjà existante pour ForCEEPS.

2.2.1 Variables de contrôle

Les scénarios de gestion généralement utilisés sont basés sur trois variables de contrôle. D'abord, la période θ définit le temps entre deux interventions. Sa valeur est en année $\theta \in [5, 10, 15, 20]$. La seconde est le type d'éclairci tp . Il influence quels arbres seront coupés : si $tp = 0$, ce seront les arbres avec le plus grand diamètre en priorité, si $tp = 1$, ce seront ceux avec le plus petit et si $tp = 0.5$, les arbres coupés seront choisis aléatoirement. Pour finir, il faut choisir la surface terrière qui restera après l'intervention G_{obj} . C'est un objectif fixe en $m^2 ha^{-1}$ ou une proportion comprise entre 0 et 1.

Dans la pratique, les interventions n'ont pas lieu tous les ans. Il faut donc définir l'ensemble des années où il y aura un contrôle T_{inv} grâce aux périodes des nb_{inv} interventions. On pose θ_k le nombre d'années entre la k^{eme} intervention et celle qui la précède.

$$T_{inv} = \left\{ \sum_{k=1}^l \theta_k \mid \forall l = 1, \dots, nb_{inv} \right\} \quad (4a)$$

Les variables $v(t_{inv})$ caractérisent l'intervention qui aura lieu l'année $t_{inv} \in T_{inv}$

$$v(t_{inv}) = (tp(t_{inv}), G_{obj}(t_{inv})) \in V \quad (4b)$$

$$V = [0, 0.5, 1] \times [10 m^2 ha^{-1}, 15 m^2 ha^{-1}, 20 m^2 ha^{-1}, 25 m^2 ha^{-1}, 30 m^2 ha^{-1}, 0.9, 0.8, 0.7]$$

Pour finir, le contrôle est défini : $\forall t \in [0, T]$

$$\begin{aligned} u(.) &= (\theta_k, v(.)) \in ([5, 10, 15, 20] \times V)^{nb_{inv}} = U \\ u(t) &= \begin{cases} v(t) = tp(t), G_{obj}(t) & \text{si } t \in T_{inv} \\ \emptyset & \text{sinon} \end{cases} \end{aligned} \quad (4)$$

2.2.2 Interventions

Maintenant que les variables de contrôles sont définies, il faut caractériser leur effet sur le système. Comme il a été dit plus tôt nous nous intéressons aux interventions de coupe. Quatre conditions permettent de caractériser l'ensemble $C(t)$ des arbres coupés l'année t . Premièrement, il faut que ce soit une année d'intervention $t \in T_{inv}$ sinon aucun arbre ne sera coupé et donc, $C(t) = \emptyset$.

Ensuite, les gestionnaires n'interviennent pas si la production est en dessous d'un seuil de surface terrière $G_{\min} = 3 \text{ m}^2 \text{ ha}^{-1}$. Cela garantit que la quantité de bois coupés compensera les dépenses faites pour l'intervention.

$$G_{\text{obj}}(t) \leq G(t) - G_{\min}$$

Notons que si la surface terrière objectif $G_{\text{obj}}(t)$ est exprimée en pourcentage, il est nécessaire de la convertir en surface terrière : $G_{\text{obj}}(t) = G_{\text{obj}}(t) \times G(t)/100$.

Si la coupe a effectivement lieu, il faut choisir les arbres. Pour classer les individus i , un score $\alpha_i(t)$ leur est attribué. Si $tp = 0.5$, par définition les arbres ont tous la même probabilité d'être coupé, le score suit donc une loi aléatoire uniforme représenté par la variable k_{α_i} , voire figure 3.

$$\alpha_i(t) = \begin{cases} (1 + C_{\max}(t) - C_i(t)) / (1 + C_{\max}(t) - C_{\min}(t)) & \text{si } tp = 0 \\ (1 + C_i(t) - C_{\min}(t)) / (1 + C_{\max}(t) - C_{\min}(t)) & \text{si } tp = 1 \\ k_{\alpha_i} & \text{si } tp = 0.5 \end{cases} \quad (5a)$$

Lors de l'intervention, les arbres sont coupés dans l'ordre croissant des scores, $\alpha_i(t)$ est donc

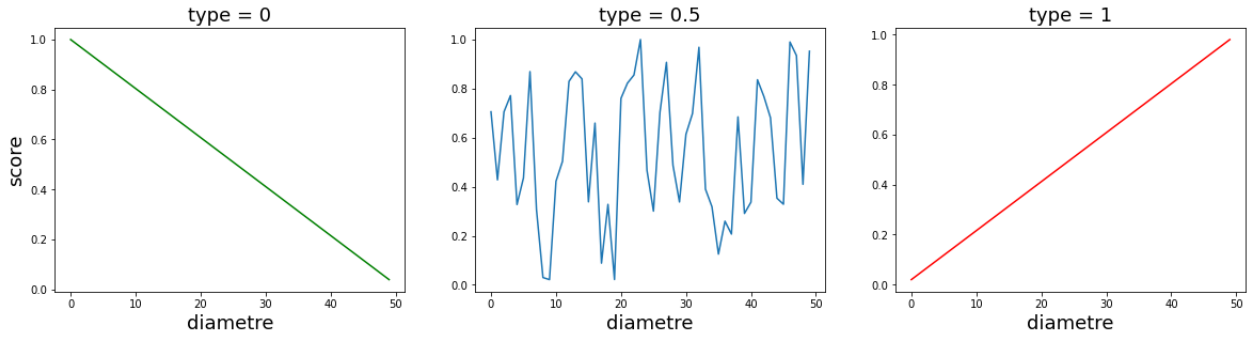


FIGURE 3 – Score α_i en fonction du diamètre D_i pour une éclaircie par le haut $tp = 0$ (gauche), uniforme $tp = 0.5$ (milieu) et par le bas $tp = 1$ (droite)

inversement proportionnel à la probabilité d'être coupé.

$$\forall i \in C(t), \forall j \in \text{pop} \setminus C(t), \alpha_i(t) \leq \alpha_j(t) \quad (5b)$$

Les arbres sont coupés jusqu'à ce que la surface terrière restante passe en dessous de l'objectif $G_{\text{obj}}(t)$. Il est donc théoriquement possible que la surface terrière $G(t)$ restante au temps t puisse être plus basse qu'espérer.

$$\begin{aligned} \sum_{i \in C(t)} G_i(t) &\leq G(t) - G_{\text{obj}}(t) \\ \exists j \in C(t) \quad \sum_{i \in C(t) \setminus \{j\}} G_i(t) &> G(t) - G_{\text{obj}}(t) \end{aligned} \quad (5c)$$

Pour finir, l'ensemble des arbres coupés $C(t)$ l'année $t \in [0, T]$ est défini :

$$C(t) = \begin{cases} \{i \in \text{pop}(t) \mid \text{5b et 5c}\} & \text{si } t \in T_{\text{inv}} \text{ et } G_{\text{obj}}(t) > G(t) - G_{\min} \\ \emptyset & \text{sinon} \end{cases} \quad (5d)$$

Il influence la dynamique du système en l'évolution de la population modélisé dans (1a) au même titre que les arbres établis $E(t)$ et ceux morts naturellement $M(t)$.

$$\text{pop}(t+1) = E(t) \cup \text{pop}(t) \setminus (M(t) \cup C(t)) \quad (5)$$

2.3 Contraintes : les objectifs écologiques et économiques

Maintenant que le contrôle (4) et son effet sur la dynamique du système (5) sont définis, il est nécessaire de savoir lesquels sont envisageables et lesquels ne le sont pas. Or, une gestion est envisageable si tous les états auxquels la dynamique du système mène sont satisfaisants. Il faut donc définir les critères de satisfaisabilité de notre système.

Pour cela, il faut analyser quels sont les services écologiques et économiques attendus de la forêt afin d'en extraire un ensemble de contraintes généralement traduit par des bornes ou des seuils. Par exemple, dans le cadre de la gestion sylvicole en forêt mélangée, les gestionnaires veulent un minimum de production tout en gardant la diversité des espèces au-dessus d'un certain seuil. Ces bornes permettent de définir un sous-ensemble de l'espace d'état qu'y est appelé espace des contraintes $K \subset X$. Ainsi, un état qui se trouve dans cet ensemble $x(t) \in K$ respecte les contraintes et est considéré comme satisfaisant.

Dans le cas étudié ici, nous avons identifié cinq critères d'intérêts. Premièrement, pour approvisionner la filière économique, il est nécessaire d'avoir au minimum de production. D'après les données de l'IGN (Institut Géographique National) [13] le critère est de $31 \text{ m}^3 \text{ ha}^{-1}$ de bois prélevé par an en moyenne. Lors de la simulation, la placette étudiée a une surface Sp qui fait généralement un dixième d'hectare, soit 1000 m^2 et ne comporte qu'un stade de développement. Pour avoir un résultat plus conforme à la réalité où les forêts font plusieurs hectares et regroupent des zones à différents stades de développement, nous avons décidé de regarder la production totale Bp en m^3 sur les T années de la simulation. Pour convertir la surface des m^2 en hectares, il faut diviser Sp par 10 000.

$$Bp = \sum_{t=1}^T \sum_{i \in C(t)} B_i(t) > 31 \times \frac{Sp}{10\,000} \times T \quad (6a)$$

avec $C(t)$ (5d) l'ensemble des arbres coupés, $B_i(t)$ (1b) leur biomasse de bois, T le nombre d'années de simulation et Sp la taille de la placette étudiée.

Ensuite, les gestionnaires veulent maintenir une certaine densité d'arbre. La métrique usuellement utilisée est la surface terrière $G(t)$ (1d) qui correspond à la somme de la surface de la section à hauteur de poitrine, soit environ à 1,30 mètre du sol, de chaque arbre. Dans ForCEEPS, elle est calculée à partir des diamètres des arbres $D_i(t)$ et d'une surface utile $Su = 20 \text{ m}^2$. D'après l'IGN une forêt a, au minimum, une surface terrière de $10 \text{ m}^2 \text{ ha}^{-1}$.

$$\forall t = 1, \dots, T \quad G(t) = \frac{\pi}{4Su} \sum_{\forall i \in \text{pop}(t)} D_i(t)^2 \geq 10 \times \frac{Sp}{10\,000} \quad (6b)$$

Pour les trois derniers objectifs, nous moyennons les valeurs sur des périodes de 5 ans qui correspondent à l'espacement minimum entre deux contrôles. Ce choix est motivé par le fait que, contrairement à la surface terrière qui ne doit jamais passer en dessous de $10 \text{ m}^2 \text{ ha}^{-1}$, les autres contraintes sont un peu plus souples. Elles seront vérifiées $\forall t_c = 5, 10, \dots, T$.

La première d'entre elles permet d'assurer une diversité structurelle. Pour caractériser cela, nous utilisons le coefficient de Gini des diamètres. Il est compris entre 0 et 1 et plus il est élevé plus l'inégalité est forte. Autrement dit, si tous les arbres ont le même diamètre, il est égale à 0 et, au contraire, il est égale à 1 si un seul individu est très différent des autres. Ici,

nous voulons qu'il soit compris entre 0.25 et 0.75 pour maximiser l'hétérogénéité structurelle.

$$m_{Gini}(t_c) = \frac{1}{5} \sum_{t=t_c-5}^{t_c} Gini(t) \in [0.25, 0.75]$$

$$Gini(t) = \frac{1}{2nb(t)} \frac{\sum_{\forall i \in \text{pop}(t)} \sum_{\forall j \in \text{pop}(t)} |D_i(t) - D_j(t)|}{\sum_{\forall k \in \text{pop}(t)} D_k(t)} \quad (6c)$$

Avec $nb(t) = |\text{pop}(t)|$ le nombre d'arbres total au début de l'année t .

Dans une forêt mélangée, les gestionnaires cherchent à maintenir la richesse de biodiversité. Il existe un certain nombre de critères pour caractériser cela. Nous choisissons un critère simple qui est la présence d'au moins deux espèces arborées sur la placette en permanence.

$$m_{nb_{\text{Esp}}}(t_c) = \frac{1}{5} \sum_{t=t_c-5}^{t_c} nb_{\text{Esp}}(t) \geq 2$$

$$nb_{\text{Esp}}(t) = \sum_{s \in \text{Esp}(t)} \begin{cases} 1 & \text{si } nb_s(t) > 0 \\ 0 & \text{sinon.} \end{cases} \quad (6d)$$

Avec $nb_s(t)$ le nombre d'arbres de l'espèce s et Esp l'ensemble des espèces pouvant être présentes dans la simulation.

Pour finir, nous voulons caractériser la bonne santé de l'écosystème. La mortalité annuelle ne doit pas excéder 25% en termes d'individu. Autrement dit le rapport entre les arbres morts et ceux toujours vivant $\tau_M(t)$ ne doit pas excéder 0.25.

$$m_{\tau_M}(t_c) = \frac{1}{5} \sum_{t=t_c-5}^{t_c} \tau_M(t) \leq 0.25$$

$$\tau_M(t) = \frac{|M(t)|}{|M(t) + \text{pop}(t)|} \quad (6e)$$

Avec $M(t)$ l'ensemble des arbres morts, $\text{pop}(t)$ (1a) l'ensemble des arbres toujours vivants.

On a maintenant défini tous les critères qui caractérisent notre espace des contraintes.

$$K = \{x = \{D_i, cs_i, \forall i \in \text{pop}\} \mid 6a, 6b, 6c, 6d, 6e\} \quad (6)$$

Tous les éléments du problème de contrôle sont définis. Nous nous intéressons maintenant à trouver des itinéraires de gestion permettant de maintenir le système dans les contraintes choisies. Après une explication de la solution adoptée, je présenterais notre cas d'application et les résultats obtenus.

3 Méthode de résolution du problème de gestion adapté aux contraintes

La comparaison entre la définition du problème de contrôle et la complexité du modèle de dynamique forestière nous ont permis de définir le type de méthode de résolution algorithmique parmi celles existantes.

L'espace de contrôle décrit dans la partie 2.2 représente beaucoup de combinaisons. Pour chacune des interventions, le type de coupe peut prendre 3 valeurs et la surface terrière

objectif peut en prendre 8, ce qui fait 24 possibilités. La période peut aussi varier entre 4 valeurs ce qui influence le nombre d'interventions durant la simulation. Si la période étudiée est de $T = 80$ ans, il y aura entre 4 et 16 interventions selon la période qui peut varier de 20 à 5 ans. L'ordre de grandeur du nombre de possibilités dépasse largement 10^{22} . Il n'est donc pas envisageable de parcourir l'ensemble de l'espace de contrôle.

Au lieu de chercher à identifier de manière exhaustive les contrôles viables V_K (2), nous explorons partiellement l'ensemble des possibilités pour identifier un sous-ensemble des politiques d'action respectant les contraintes que nous avons défini (6). L'analyse du sous-ensemble des contrôles identifiés permettra peut-être de trouver des schémas de gestion permettant de maintenir les services écologiques et économiques de la forêt dans le temps.

Afin de faire cela, nous considérons notre problème comme un problème d'optimisation en grande dimension. Les contrôles possibles ne sont plus classés de manière binaire en fonction de s'ils permettent le respect des contraintes ou non, mais proportionnellement à leur capacité à résoudre le problème. Nous utilisons des heuristiques pour parcourir l'espace des contrôles afin de trouver un sous-ensemble se rapprochant le plus possible de l'ensemble des contrôles viables. Pour analyser cet ensemble des résultats nous utilisons une méthode de classification sur facteurs.

3.1 Résolution par algorithme génétique

La solution que nous choisissons est de la famille des algorithmes évolutionnaires. Cette famille d'algorithmes d'optimisation est basée sur la simulation du processus d'évolution naturelle. Le principe général est d'étudier une population dont chaque individu est une solution au problème d'intérêt. Des processus aléatoires sont utilisés itérativement pour faire évoluer cette population. À chaque itération, la nouvelle génération est composée d'individus correspondant à de meilleures solutions au problème. Pour utiliser ces algorithmes, il est nécessaire de définir une métrique permettant de mesurer l'efficacité des individus à résoudre le problème. C'est la fonction objectif qui joue ce rôle en attribuant à chaque individu un score appelé aussi fitness. Généralement plus le score d'individu est grand plus c'est une solution efficace au problème étudié.

Nous nous intéressons plus précisément aux algorithmes génétiques (GA pour Genetic Algorithms). Ils sont inspirés de l'évolution génétique des populations des êtres vivants. Chaque individu a un génome, représenté par un vecteur de bits ou de réels. Chacun des éléments de ce vecteur est appelé gène. Le génome représente une solution au problème, c'est donc lui qui permet de calculer le score de l'individu. Ce score sert à déterminer la survie de l'individu tout au long de l'algorithme. L'évolution de la population se fait grâce à la succession d'événements tels que le brassage génétique ou les mutations, au sens de l'algorithme génétique. Les GA ont déjà été utilisés pour résoudre des problèmes issus de la théorie du contrôle [14, 15, 16].

3.1.1 Description générale de l'algorithme génétique

L'algorithme génétique que j'ai utilisé lors de ce stage est illustré dans la figure 4. La première étape est d'initialiser une population aléatoirement de n_{init} individus.

La deuxième étape consiste à appliquer itérativement plusieurs étapes. D'abord, il attribue un score à chaque individu de la génération via la fonction objectif détaillée dans la suite. Plus ce score est grand plus l'individu est jugé efficace et, au contraire, un individu avec un score

Algorithme principal

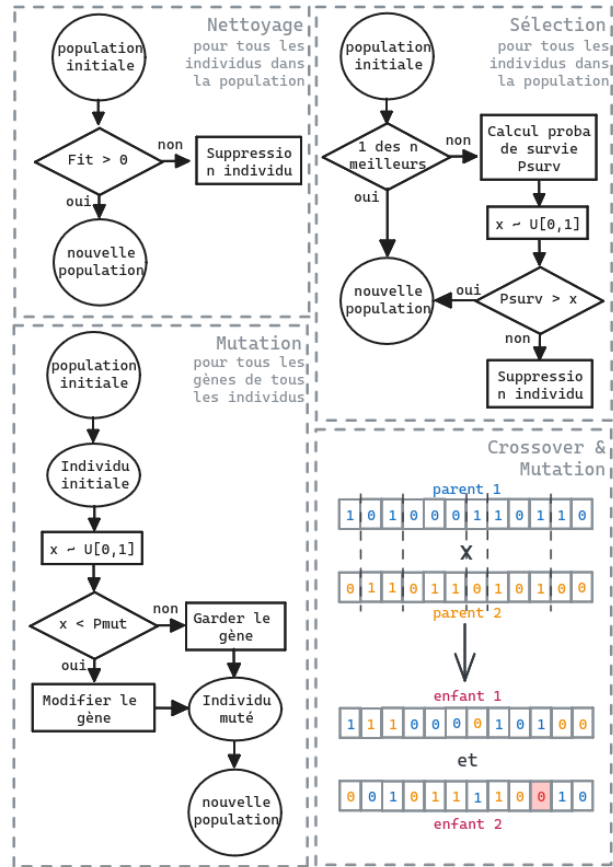
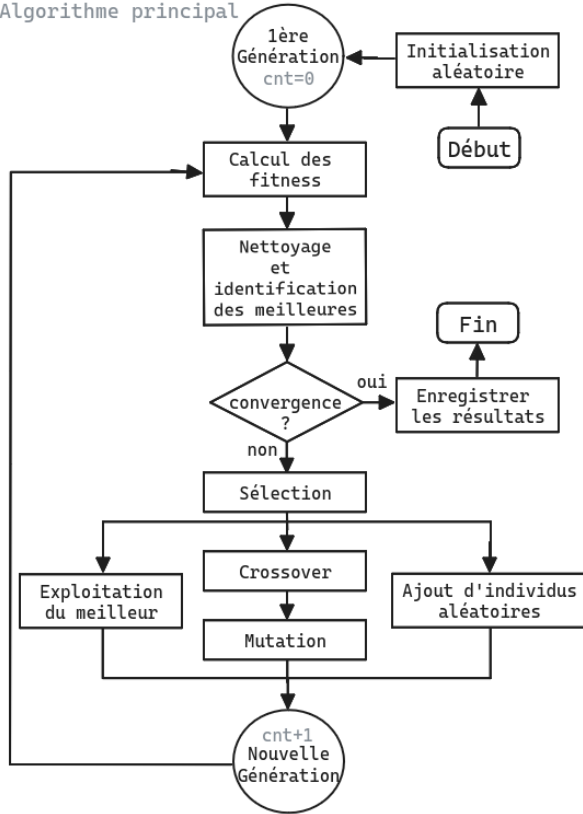


FIGURE 4 – Représentation du fonctionnement général de l’algorithme génétique (à gauche). Détail des étapes de Nettoyage, de Sélection et de Mutation (à droite). Illustration du brassage génétique (en bas à droite), les gènes issus du parent 1 sont en bleus, ceux du parent 2 en orange, les points de croisement sont en pointillés et le gène muté est en rouge.

de 0 n’est pas apte à résoudre le problème. L’évaluation permet de nettoyer la population, en retirant tous les individus avec un score nul et d’identifier les meilleurs de la génération. Si le critère d’arrêt n’est pas atteint, l’algorithme applique une opération de sélection dite Darwinienne, voir la figure 4. Une probabilité de survie proportionnelle au score est associée aux individus. Cette probabilité est comparée à la réalisation d’une variable aléatoire uniforme pour savoir si les individus survivent. Lors de cette étape l’algorithme fixe à 1 la probabilité de survie des n_{save} meilleurs pour leur garantir de passer cette étape.

La troisième étapes, correspond au calcul de la nouvelle génération. Cela se fait en trois sous-étapes. D’abord, le brassage génétique ou crossover permet d’explorer l’espace en créant de nouveaux individus. Des paires de parents sont sélectionnés aléatoirement et leurs gènes sont mélangés pour créer deux nouveaux individus appelés enfants. Lors de cette sous-étape, il est possible de voir des gènes changer de valeurs. c’est ce qu’y est appelée une mutation de gènes et elles apparaissent avec une probabilité de P_{mut} . Dans la figure 4 (en bas à droite) les gènes des parents sont échangés au niveau des points de croisement et le génome des enfants est un mélange des gènes des parents 1 et 2. Le gène de l’enfant 2 marqué en rouge a muté, il aurait dû être égale à 1, comme pour celui du parent 2, mais il est égal à 0. Ensuite vient la deuxième sous-étape, l’exploitation du meilleur. Cela consiste à appliquer alternativement des mutations à chacun des gènes du meilleur individu. Le meilleur dans l’ensemble de l’individu

et de ses mutations est gardé dans la prochaine génération. Ce processus permet une recherche locale autour du meilleur individu. Enfin, la dernière sous-étape vise à augmenter la diversité et mieux explorer l'espace des possibles, j'ajoute n_{rand} individus aléatoires.

L'algorithme repart de l'étape de l'évaluation des individus et continu itérativement jusqu'à ce que le critère soit atteint. Dans notre cas, j'ai utilisé un critère de stagnation du meilleur score. Ce critère consiste à définir que, si le meilleur score n'évolue pas durant cinq générations, alors l'optimum est atteint et l'algorithme s'arrête. L'algorithme sauvegarde alors les résultats dans différents fichiers de sortie.

Il est nécessaire, pour utiliser cette méthode, de convertir notre problème de gestion forestière en problème génétique. Pour rappel, le problème que nous étudions lors de ce stage est le suivant : en partant d'un état initial de la forêt $x_0 = \{D_i(0), cs_i(0), \forall i \in \text{pop}(0)\} \in K$, nous cherchons l'ensemble $V_K(x_0)$ (2) qui contient au moins un itinéraire de gestion $u \in U$, permettant de maintenir la forêt dans les contraintes K que nous avons fixées. Il nous faut convertir ce problème pour appliquer l'algorithme génétique. Dans la suite, je décrirai comment sont définis les individus et ce qu'ils représentent. Puis quelques particularités des opérations de brassage génétique et de mutation seront expliquées. Enfin, je discuterai de la fonction objectif choisie.

3.1.2 Définition des individus de l'algorithme génétique

Dans les algorithmes génétiques, les individus représentent des solutions aux problèmes. Dans notre cas nous nous intéressons au contrôle $u \in U$ (4) qui définit l'ensemble des interventions de la simulation. Une intervention est caractérisée par trois variables : la période qui la sépare de la dernière intervention θ , le type de coupe tp et la surface terrière objective G_{obj} .

Pour utiliser l'algorithme il faut associer à chaque individu j un génome Y^j . Chaque intervention est représentée par un ensemble de trois gènes. L'itinéraire de gestion qui en résulte est représenté par l'ensemble de ces gènes regroupés dans un seul vecteur de réel. Dans un algorithme génétique il est nécessaire que chaque génome soit de la même longueur, or le nombre d'interventions n'est pas fixe d'un contrôle à l'autre. Par la suite, la durée de simulation est fixée à $T = 80$ ans. Il y aura au plus une intervention tous les 5 ans, soit 16 au total. Un génome Y^j est donc composé de $16 \times 3 = 48$ réels. Il peut être défini comme suit :

$$\begin{aligned} Y^j &= (\theta_k^j, tp_k^j, G_{\text{obj}k}^j)_{k=1, \dots, 16} \\ Y^j &\in ([0, 5, 10, 15, 20] \times [0, 0.5, 1] \times [0, 10, 15, 20, 25, 30, 0.9, 0.8, 0.7])^{16} \subset \mathbb{R}^{48} \end{aligned} \quad (7a)$$

Ainsi, pour toutes les années $t = 5, 10, \dots, 80$, l'intervention est représentée par un trio de gènes $k = 1, 2, \dots, 16$. Dans le cas où, pour une année t , $\theta_k^j = 10$ cela signifie qu'il n'y a pas de coupe durant dix ans et donc pas d'intervention lors de l'année $t - 5$, je fixe donc $(\theta_{k-1}^j, tp_{k-1}^j, G_{\text{obj}k-1}^j) = (0, 0, 0)$. De même, si $\theta_k^j = 15$ ou 20 il y a six ou neuf zéros pour signifier qu'il n'y a pas de coupe les années précédentes.

Ces zéros doivent être considérés comme des gènes à part entière lors des opérations d'évolution de la population. La procédure est détaillée dans la partie qui suit.

3.1.3 Opérations réalisées dans l'algorithme génétique

Les zéros associés aux périodes entre les interventions supérieures à 5 ans ne peuvent pas être traités comme des gènes classiques. En effet, un gène $\theta_k^j > 5$ ne peut pas être séparé des

zéros qui le précèdent. Cela a une influence sur deux des opérations d'évolution : le brassage génétique et la mutation

3.1.3.1 Brassage génétique

Les points de croisement, représentés en pointillés sur la figure 4, doivent être définis lors du brassage. Ici, il y a un point de croisement entre chaque gène, sauf s'il y a des zéros représentant une non-intervention chez un des deux parents. Dans ce cas les zéros et le θ_k^j restent ensemble chez un des deux enfants, et par conséquent les gènes correspondants du second parent aussi. Un exemple est donné dans la figure 5.

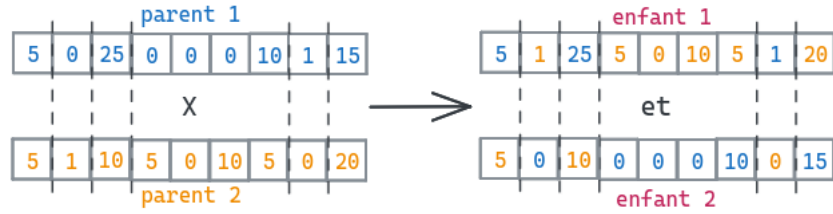


FIGURE 5 – Exemple d'un brassage génétique (crossover) de l'algorithme génétique utilisé. Les gènes issus du parent 1 sont en bleus, ceux du parent 2 en orange et les points de croisement sont en pointillés.

3.1.3.2 Mutation

Les mutations peuvent survenir aléatoirement lors du brassage génétique (voir figure 4) ou systématiquement lors de l'exploration du meilleur. Théoriquement, elles peuvent avoir lieu sur n'importe quel gène. Cependant, ici, j'ai fait le choix de ne jamais faire muter les gènes correspondant au contrôle sur la période θ_k^j . Cela est motivé par le fait qu'une mutation sur ce gène entraîne l'ajout ou la suppression des zéros de non-intervention. Si on diminue la période, il faut alors choisir quelles valeurs remplaceront les trois gènes précédemment égaux à zéro. De plus, la modification des périodes influe grandement sur le contrôle en faisant varier le nombre d'interventions.

Pour préserver l'identité de l'individu et éviter le choix partiel des gènes ajoutés, j'ai préféré, pour l'instant, ne faire muter que les gènes correspondant au type de coupe tp_k^j et à la surface terrière objective $G_{obj_k}^j$. La mutation a lieu avec une probabilité P_{mut} . Le gène prend alors soit la valeur au-dessus, soit celle en dessous avec autant de probabilité. Si le gène a déjà une valeur extrême, alors il prendra la valeur la plus proche. Par exemple, le type de coupe peut prendre trois valeurs : $[0, 0.5, 1]$, si $tp_k^j = 0.5$ alors il peut muter en 0 ou en 1, mais si $tp_k^j = 0$ ou si $tp_k^j = 1$ alors il mute automatiquement en 0.5.

Maintenant que les opérations sont détaillées, il ne manque plus qu'à définir comment sélectionner les individus. Pour cela il est nécessaire d'avoir un moyen de les classer en fonction de leur capacité à résoudre le problème.

3.1.4 Fonction objectif

Lors de ce stage, nous voulions savoir si les itinéraires de gestion envisagés permettaient d'assurer les services forestiers définis plus tôt par l'ensemble des contraintes $K(6)$ et donc, la fonction objectif doit évaluer la capacité d'un individu à résoudre le problème. Ici, nous

voulons qu'elle reflète si le contrôle, correspondant à l'individu, permet de garder le système dans les contraintes K .

Pour cela, la métrique que j'ai utilisée correspond à la somme des distances aux contraintes tout au long de la simulation. Les contraintes n'ont pas tous la même dimension temporelle, elles n'apparaissent donc pas le même nombre de fois dans la somme. Celle sur la production n'est calculée qu'une fois pour l'entièreté de la simulation. Celle sur la surface terrière 80 fois car c'est une contrainte annuelle. Les trois autres sont moyennées sur 5 ans et ont donc 16 occurrences dans la somme. Il faut donc normaliser en fonction de ces apparitions. De plus, les grandeurs ne sont pas les mêmes, pour maintenir une équité entre les différentes contraintes dans la somme, les distances sont divisées par leur écart-type. J'ai défini pour l'individu j dont le génome Y^j décrit le contrôle u^j :

$$\Phi(u^j) = \begin{cases} \phi(u^j) & \text{si } u^j \in V_K(x_0) \\ 0 & \text{sinon.} \end{cases} \quad (7b)$$

$$\phi(u^j) = \frac{Bp - 31}{Bp_{\max} - 31} + \frac{1}{80} \sum_{t=1}^{80} \frac{G(t) - 10}{G_{\max} - 10} + \frac{1}{16} \sum_{t_c=1}^{16} \left(\frac{|m_{Gini}(t_c) - 0.5|}{0.5} + \frac{m_{nbEsp}(t_c) - 2}{nbEsp_{\max} - 2} + \frac{0.25 - m_{\tau_M}(t_c)}{0.25} \right)$$

où Bp_{\max} , G_{\max} et $nbEsp_{\max}$ sont respectivement la production en $m^3 ha^{-1}$, la surface terrière en m^2 et le nombre d'espèces maximum qui peuvent être atteints durant la simulation. Pour finir, si lors de la simulation le système sort des contraintes, le score de l'individu sera considéré comme nul et il sera alors sorti de la population lors de la phase de nettoyage.

Il est important de noter que dans ForCEEPS, l'établissement et la mortalité sont des processus partiellement stochastiques. La dynamique de la forêt peut donc légèrement varier d'une simulation à une autre. Cela influence son état et par conséquent le calcul du score. Il est possible de lancer la simulation plusieurs fois simultanément grâce à des patchs. Le calcul de Φ se fait avec la valeur moyenne obtenue sur l'ensemble de ces patchs. Cependant, il existe plusieurs méthodes pour définir si le contrôle permet de rester dans les contraintes. Voici les deux qui ont été retenus.

3.1.4.1 Méthode basée sur l'aversion au risque

Cette méthode consiste à regarder le pire des cas. Ici, où x_p est l'état de la forêt calculé avec f_p la dynamique lors de la simulation avec le patch p , on a :

$$u^j \in V_K(x_0) \Leftrightarrow \forall p \forall t \ x_p(t+1) = f_p(x_p(t), u(t)) \in K$$

Autrement dit, si une seule des simulations sort des contraintes alors, l'individu a un score de 0 et est supprimé de la population. Cela permet de s'assurer de la validité des schémas sélectionnés. Mais cela limite aussi l'espace des recherches.

3.1.4.2 Méthode basée sur le calcul de la moyenne

Pour cette méthode, nous regardons si les contraintes sont respectées en moyenne sur les P patchs lancés, tel que :

$$u^j \in V_K(x_0) \Leftrightarrow \forall t \ \frac{1}{P} \sum_{p=1}^P x_p(t+1) = f_p(x_p(t), u(t)) \in K$$

Il est donc possible de garder un schéma qui sort parfois des contraintes. Ce choix moins contraignant que l’aversion au risque est motivé par le fait que FoCEEPS est un modèle et ne représente pas totalement la réalité. Lors de la mise en application, décrite dans la section 4, la version utilisée est celle de la moyenne.

3.1.5 Analyse de sensibilité de l’algorithme

Avant d’utiliser l’algorithme génétique, il est nécessaire de le régler pour parcourir le plus représentativement possible l’ensemble des possibles. Pour cela je fais une analyse de sensibilité de l’algorithme face aux perturbations de ces paramètres. Quatre paramètres ont été perturbés : n_{init} représentant le nombre d’individus générés aléatoirement lors de l’initialisation pour composer la première génération, n_{save} le nombre des meilleurs individus sauvés automatiquement lors de la sélection, P_{mut} la probabilité qu’un gène mute lors des croisements et n_{rand} le nombre d’individus aléatoire ajoutés à chaque génération. Pour chaque paramètre je teste entre deux et quatre valeurs et je lance l’algorithme pour toutes les combinaisons. Pour ce test, j’utilise l’inventaire d’intérêt et les données climatiques historiques de la ville de Berne de 1950 à 2020, complétées avec 10 ans de simulations.

Tous les résultats obtenus sont présentés en figure 6. Contrairement au problème d’optimisation classique, la valeur de la fonction objectif n’est pas le critère qui nous intéresse le plus pour choisir la meilleure combinaison de paramètre. En effet, ici nous cherchons à parcourir l’espace le plus largement et le plus efficacement possible. La variance des valeurs de la fonction objectif ainsi que le nombre d’individus viables calculés par l’algorithme permettent de se faire une idée de la diversité des solutions identifiées. Si la variance reste faible, dans notre cas inférieure à 0.03, pour toutes les configurations, le nombre d’individus, lui, est plus élevé pour $n_{\text{init}} = 500$. La proportion d’individu viable peut donner, quant à elle, une idée de l’efficacité de l’algorithme : plus elle est élevée, moins l’algorithme a créé d’individus qui sont sortis des contraintes. Pour finir, il est préférable d’avoir un nombre de générations pas trop faible pour parcourir plus largement l’espace sans tomber directement dans un optimum qui pourrait être local.

En prenant en compte ces différents critères nous choisissons par la suite la combinaison $n_{\text{init}} = 500$, $n_{\text{save}} = 5$, $P_{\text{mut}} = 0.1$ et $n_{\text{rand}} = 10$, les carrés verts sur la figure 6.

3.2 Classification des résultats

Pour analyser ces résultats nous classons les individus viables identifiés par l’algorithme génétique. J’utilise une méthode de classification à ascendance hiérarchique combinée avec une analyse en composante principale. Je classe les individus en me basant sur les cinq variables de contraintes qui permettent le calcul de la fonction objectif.

3.2.1 Analyse en composantes principales

L’analyse en composantes principales (ACP) est une méthode permettant d’identifier des proximités entre des individus caractérisés par de multiples variables en résumant l’information dans quelques variables synthétiques [17]. Pour cette analyse les variables que j’utilise sont les contraintes qui permettent de savoir si un individu est viable et de calculer la fonction objectif. Pour chaque individu, je récupère la moyenne de la production Bp , la surface terrière G , le coefficient de *Gini*, le nombre d’espèce nb_{Esp} et le taux de mortalité τ_M tout au

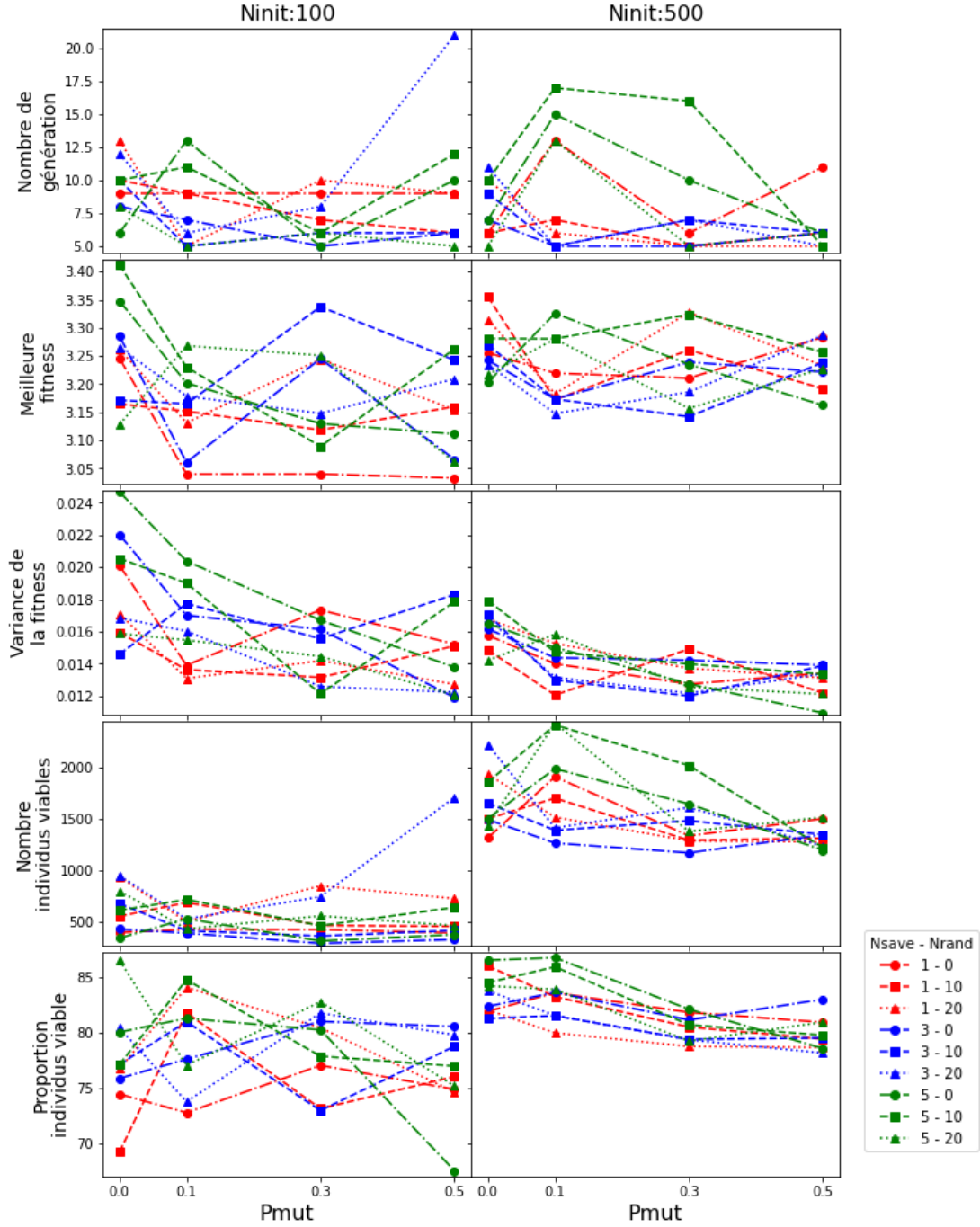


FIGURE 6 – Résultat de l’analyse de sensibilité des paramètres de l’algorithme génétique avec un climat historique. Nombre de générations, meilleur score calculé, variance des scores de l’ensemble des individus viables, nombre d’individus viables et proportion d’individus viables par rapport au nombre total d’individus calculés lors du GA en fonction de P_{mut} pour différentes configurations de n_{save} et n_{rand} et pour $n_{\text{init}} = 100$ (à droite) et $n_{\text{init}} = 500$ (à gauche).

long de la simulation. J'utilise une ACP normée pour que chaque variable ait le même poids dans l'analyse.

Il est possible de caractériser leurs liaisons en calculant des coefficients de corrélation croisés. J'utilise la corrélation de Pearson qui caractérise la relation linéaire entre deux variables continues.

Pour diminuer le nombre de variables en les synthétisant, il faut hiérarchiser la contribution des variables dans la différenciation des individus. Il est possible de le faire grâce aux valeurs propres de la matrice de corrélation, obtenues en la diagonalisant dans l'espace des variables. Chaque dimension correspond à un vecteur propre. Sa valeur propre caractérise l'importance que joue cette dimension dans la différenciation des individus. Le vecteur propre nous permet de connaître la contribution qu'a chaque variable dans la dimension.

3.2.2 Classification à ascendance hiérarchique

Les méthodes de classification ascendante hiérarchique (CAH) sont des méthodes qui n'utilisent pas les données brutes des individus mais des coordonnées factorielles. Ces coordonnées permettent de calculer les distances entre individus afin de les regrouper pour former une arborescence. La méthode est la suivante : les deux individus les plus proches sont rassemblés et remplacés par leur barycentre ainsi de suite jusqu'à ce que tous les individus soient regroupés.

J'utilise les coordonnées factorielles obtenues avec l'ACP ainsi que la distance de Ward [18]. Cette dernière est calculée, pour deux groupes CL_1 et CL_2 , comme le carré de la distance euclidienne entre leur barycentre, B_1 et B_2 , pondéré par leur effectif respectif, n_1 et n_2 .

$$D_w(CL_1, CL_2) = \|B_1 - B_2\|^2 \frac{n_1 n_2}{n_1 + n_2}$$

La méthode de Ward cherche à minimiser l'inertie intra-classe et à maximiser l'inertie inter-classe afin d'obtenir des classes les plus homogènes possibles.

La CAH permet la construction de dendrogramme (voir la figure 9). Cet arbre nous permet de choisir visuellement le nombre de classes. L'idée est de prendre un niveau de coupe (en pointillés gris) qui intercepte des branches relativement longues. Cela permet de maximiser les distances entre les individus des différentes classes. Pour plus de détail, un test de pertinence pour du nombre de classes est présenté dans l'annexe C.

Maintenant que l'ensemble des méthodes utilisées sont définies, je peux les appliquer à notre cas d'étude. Il est décrit, ainsi que les résultats, dans la suite.

4 Mise en application

4.1 Cas d'études

Notre cas d'étude est la forêt de Brun située à proximité des locaux de l'INRAE Champenoux (54113), proche de Nancy, dans l'Est de la France. C'est une forêt de six cents hectares composée de onze d'espèces arboricoles, majoritairement du chêne et du charme. Elle appartient à AgroParisTech qui s'occupe aussi de sa gestion.

Nous avons défini les données d'entrée du modèle correspondant à ce cas. La surface de la placette que nous regardons est de $Sp = 1000 \text{ m}^2$, le site se trouve à une latitude $Lat = 43.93$.

La quantité de nutriment dans le sol est de $N_{\text{soil}} = 100$ et la présence d’animaux correspond à un indice d’abrutissement $BrI = 0.2$. Pour finir, le point initial des simulations x_0 est défini grâce à un inventaire partiel fait par échantillonnage allégé, datant de juillet 2023.

Nous utilisons des données climatiques issues de modèle de prédiction basé sur les scénarios du rapport du GIEC de 2014 [10]. Notre choix s’est porté sur les scénarios de prédiction à plus 4.5 et 8.5 W m^{-2} , RCP 4.5 et 8.5, de deux modèles : celui du CNRM (Centre National de Recherche Météorologique) et celui du MPI (Max-Planck-Institut für meteorologie). Ils ont tous les deux été testés et comparés à différents modèles dans plusieurs études [19, 20, 21]. Pour l’Ouest européen, le modèle du MPI est considéré comme dans la moyenne alors que celui du CNRM est plus clément, notamment concernant les précipitations.

Pour finir, j’ai lancé des simulations pour estimer la production Bp_{max} , la surface terrière G_{max} maximum afin de calculer la fonction objectif. Avec $Bp_{\text{max}} = 125 \text{ m}^3 \text{ ha}^{-1}$, $G_{\text{max}} = 20 \text{ m}^2$ et le nombre d’espèce maximum $nb_{\text{Esp}_{\text{max}}} = 11$, (7b) devient :

$$\phi(u_j) = \frac{Bp - 31}{94} + \frac{1}{80} \sum_{t=1}^{80} \frac{G(t) - 10}{15} + \frac{1}{16} \sum_{t_c=1}^{16} \left(\frac{|m_{Gini}(t_c) - 0.5|}{0.5} + \frac{m_{nb_{\text{Esp}}}(t_c) - 2}{9} + \frac{0.25 - m_{\tau_M}(t_c)}{0.25} \right)$$

4.2 Présentation des résultats de l’algorithme

Je fais tourner l’algorithme en partant de l’inventaire de la forêt de Brun avec les paramètres choisis plus tôt ($n_{\text{init}} = 500$, $n_{\text{save}} = 5$, $P_{\text{mut}} = 0.1$ et $n_{\text{rand}} = 10$). Pour les quatre jeux de données, je le lance cinq fois l’algorithme. Les résultats obtenus sont présentés dans le tableau 1.

| | | Nombre de génération | Score max | Variance score | Nombre d’in- dividu viable | % d’individu viable |
|-----|------|-------------------------|-----------|----------------------|-------------------------------|------------------------|
| rcp | MPI | 9 | 2.89 | 2.4×10^{-2} | 1664 | 82.5 |
| 4.5 | CNRM | 9 | 2.94 | 2.5×10^{-2} | 1720 | 81.3 |
| rcp | MPI | 9 | 2.20 | 4.1×10^{-3} | 2120 | 82.8 |
| 8.5 | CNRM | 7 | 2.22 | 4.6×10^{-3} | 1683 | 82.5 |

TABLE 1 – Résultats moyennés pour cinq lancements de l’algorithme génétique avec l’inventaire de la forêt de Brun, les données climatiques des modèles du MPI et du CNRM suivant les RCP 4.5 et 8.5 pour la configuration de paramètre $n_{\text{init}} = 500$, $n_{\text{save}} = 5$, $P_{\text{mut}} = 0.1$ et $n_{\text{rand}} = 10$.

4.2.1 Analyse en composantes principales

La première étape de l’ACP est de caractériser les liens entre les variables de contrôle. Les résultats obtenus sont présentés dans la figure 7. Il y a plusieurs remarques à faire sur ces graphiques. D’abord, les résultats sont très ressemblants avec les quatre jeux de données climatiques. La corrélation des contraintes est donc sûrement inhérente au modèle ForCEEPS et à l’algorithme génétique mais ne dépend pas des données climatiques utilisées. Ensuite, quatre variables sont corrélées positivement deux à deux : le nombre d’espèces avec la production et le taux de mortalité avec la surface terrière. De plus, ces deux groupes sont fortement corrélés négativement entre eux. Le coefficient de Gini, lui, ne semble pas corrélé avec aux autres variables.

L’analyse de l’ACP, représenté en figure 8, montre que les deux premières composantes contribuent pour respectivement 57% et 26% de la variance totale, ce qui montre que deux

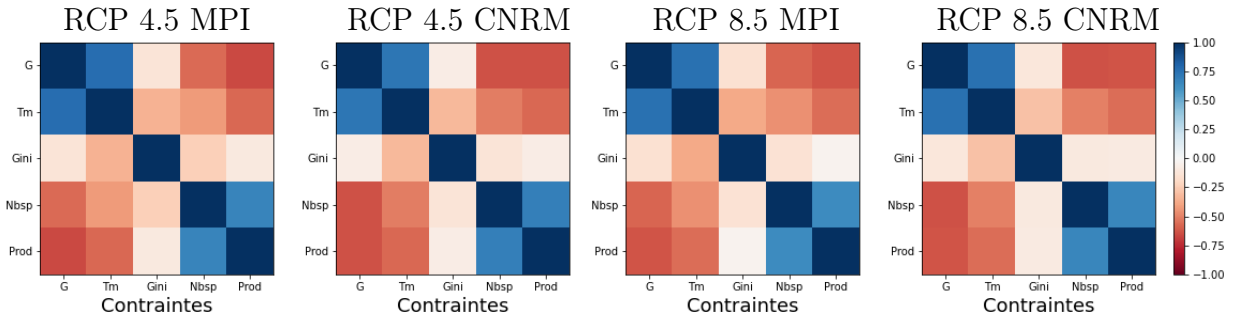


FIGURE 7 – Corrélation de Pearson des contraintes d’après l’ACP sur l’ensemble des individus viables identifiés par le GA avec les données climatiques des modèles du MPI et du CNRM suivant les RCP 4.5 et 8.5. L’échelle de couleur représente le niveau de corrélation : bleu foncé fortement corrélé positivement, blanc non corrélé et rouge foncé fortement corrélé négativement. Dans la suite, si les résultats ne sont présentés que pour un seul jeu de données, les figures correspondante pour les autres jeux de données sont en annexe D.

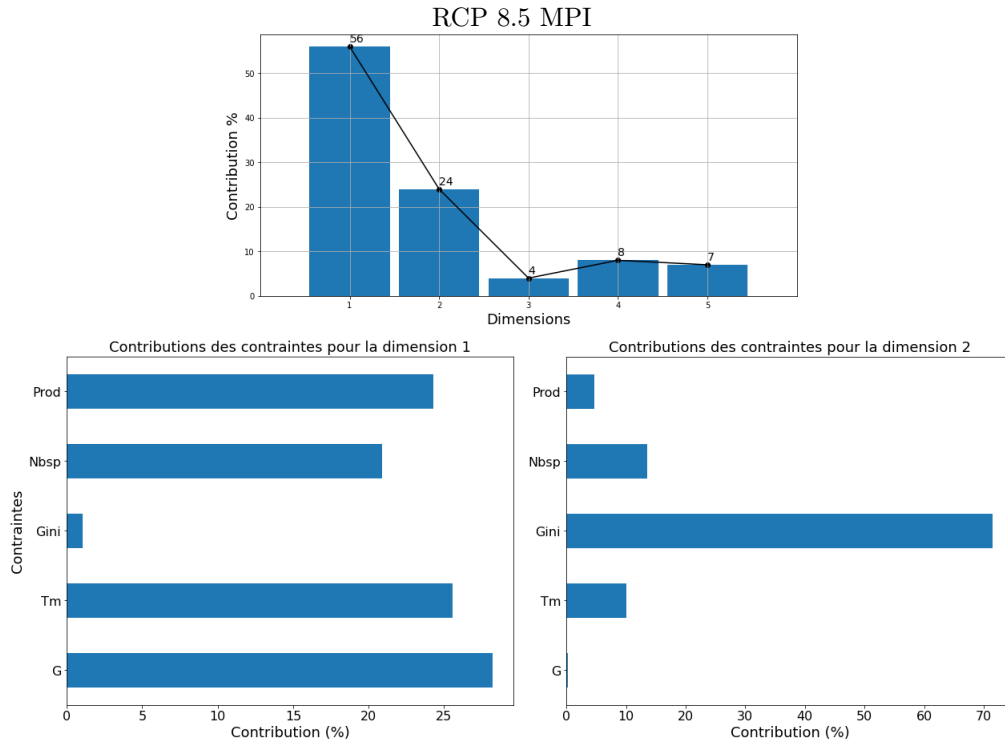


FIGURE 8 – Contribution des dimensions (en haut) calculées avec l’ACP sur les contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant le RCP 8.5. Contribution des contraintes pour les deux dimensions principales (en bas). Figures complémentaires 18, 19.

axes représentent la variabilité de notre jeu de données. La deuxième dimension représente principalement le coefficient de Gini, alors que la première représente à peu près à part égale la contribution des quatre autres variables. Ce résultat est cohérent avec les matrices de corrélation : le coefficient de Gini n'est corrélé à aucune des autres variables qui, elles, sont fortement corrélées positivement ou négativement entre elles. Deux dimensions suffisent donc à expliquer la plupart des informations représentées par les cinq variables de contrainte.

Il est possible de représenter graphiquement les individus, comme cela est fait dans la figure 9 (à droite), en fonction des composantes principales. La projection des vecteurs de contraintes permet d'évaluer approximativement les variables qui lui sont associées. Par exemple, si un individu est en bas à droite du graphique, cela signifie qu'il a une production et un nombre d'espèces élevé comparativement aux autres car c'est dans cette direction que sont les flèches représentant ces variables. Au contraire, ses autres contraintes doivent avoir une valeur relativement basse.

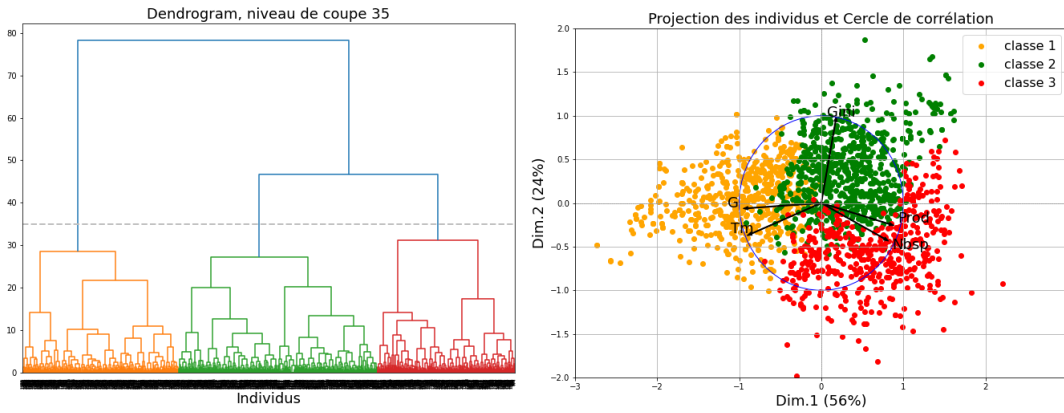


FIGURE 9 – Résultats de la CAH avec la distance de Ward sur les données de l'ACP sur l'ensemble des contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant le RCP 4.5. Le dendrogramme (à gauche) représente les individus en abscisse et leurs distances en ordonnée. La projection des individus et du cercle de corrélation des contraintes sur les deux composantes principales (à droite). Les couleurs correspondent aux classes d'individus, les pointillés gris sont le niveau de coupe du dendrogramme, les flèches représentent les coefficients de corrélation entre les variables et les deux composantes principales. Figure complémentaire 20.

4.2.2 Classification à ascendance hiérarchique

Le dendrogramme obtenu après la CAH est présenté à gauche de la figure 9. Pour les résultats obtenus avec les données RCP 8.5 du modèle MPI, nous avons fait le choix de prendre trois classes. Les projections, à droite de la figure 9 associées aux représentations des contraintes dans l'espace des composantes principales, permettent de supposer les caractéristiques communes aux individus d'une même classe. Ici, les individus de la première classe (en jaune) semblent avoir une surface terrière et un taux de mortalité élevé et peu d'espèces et de production. C'est l'inverse pour ceux de la troisième classe (en rouge). Enfin, la classe deux (en vert) paraît être celle favorisant le plus le coefficient de Gini.

Il est possible d'étudier l'évolution des simulations associées aux individus pour vérifier et affiner ces interprétations. La figure 10 représente l'évolution médiane pour chaque classe ainsi que l'écart entre le premier quartile (0.25) et le troisième (0.75). Les courbes se rapprochent des hypothèses. Les individus de la première classe (en jaune) ont un nombre d'espèces et une

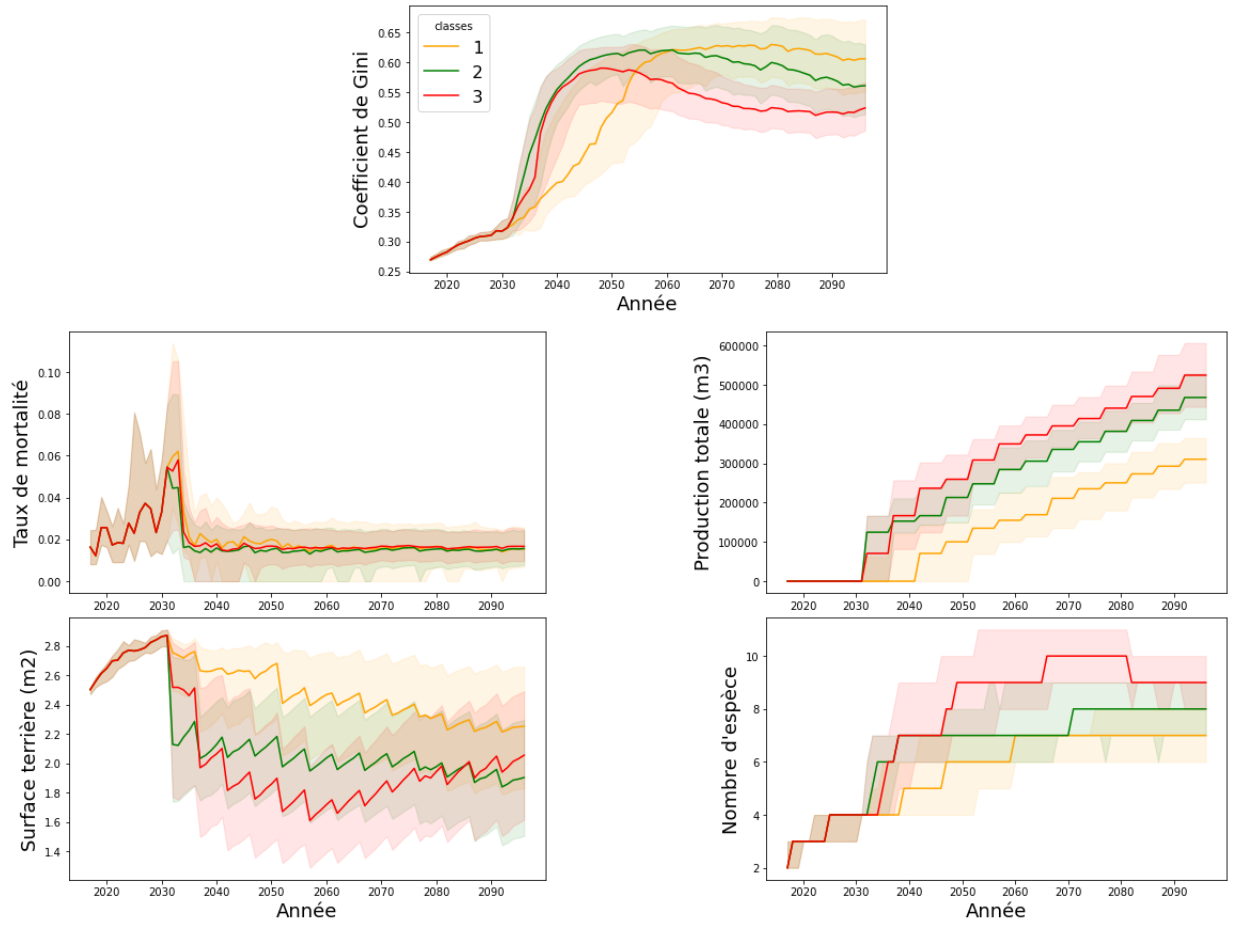


FIGURE 10 – Médiane et écart interquartile (0.25-0.75) de l'évolution au cours de la simulation des contraintes, coefficient de Gini, taux de mortalité, surface terrière, production et nombre d'espèce, pour les différentes classes obtenues lors de la CAH avec la distance de Ward sur les données de ACP sur l'ensemble des contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant le RCP 4.5. Figure complémentaire 21.

production inférieure aux individus des autres classes. Leur surface terrière est en moyenne supérieure, même si comme le montre le troisième quartile, certains individus des autres classes peuvent avoir une surface terrière plus grande sur les vingt dernières années. Les taux de mortalité ne diffèrent pas significativement selon la classe. Comme le laisse deviner la projection sur les composantes principales, la tendance la troisième classe (en rouge) est inversé à celle de la première. Enfin la médiane des individus de la classe deux (en vert) est située entre celle des deux autres pour la production, le nombre d'espèces et la surface terrière. Pour ce qui est du coefficient de Gini, en moyenne sur les 80 ans, il semble plus grand pour la classe deux.

Pour répondre à notre question initiale nous voulons identifier des itinéraires de gestion type associées à chaque classe. Afin de faire cela, nous avons étudié les itinéraires décrits par le génome de chaque individu. Comme il a été dit lors de la définition du contrôle, il arrive qu'une intervention prévue ne soit pas effectuée. C'est le cas si elle n'est pas assez productive par exemple. Nous avons comparé les interventions programmées à celle réellement effectuées dans la figure 11. Le nombre d'interventions programmées est plus grand que celui des effectués. De même, les périodes entre les interventions sont plus courtes que prévues. Cela

confirme que certaines interventions programmées ne sont pas effectuées. De plus, si le type de coupe reste globalement identique, la moyenne d'intensité de coupe est plus élevée dans la réalité que ce qui était programmé. Les interventions non-effectuées ne sont donc pas celle dont la surface terrière à coupé est la plus élevée.

Ensuite, les interventions effectuées présentent plus de variation selon les classes. Dans la classe une, il y a moins d'interventions et par conséquent les périodes qui les séparent sont plus longues. Le pourcentage de surface terrière coupée lors des interventions est le critère qui présente le plus de différence selon la classe, la médiane est respectivement égale est, en moyenne, d'autant plus grand que le rang de la classe. Pour finir, le type de coupe est majoritairement égal à 0.5 pour toutes les classes, soit une sélection homogène dans les classes de diamètre.

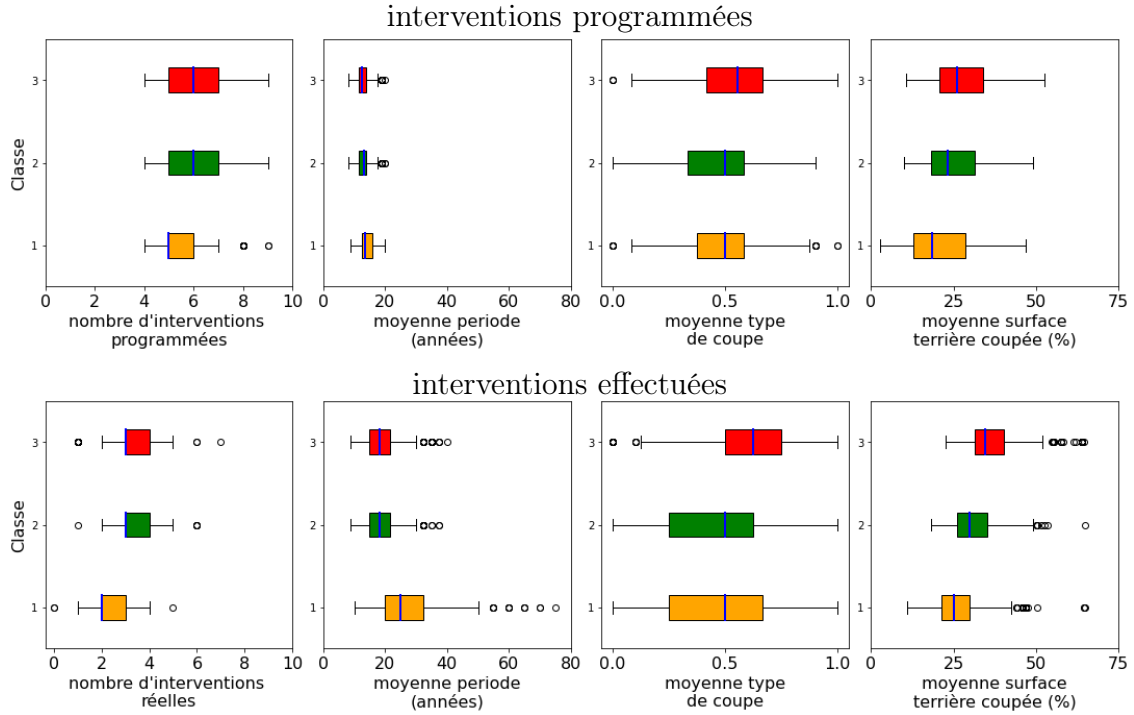


FIGURE 11 – Médiane (en bleu) et écart interquartile du nombre d'intervention et des moyennes sur la simulation de la période entre deux interventions, du type de coupe et du pourcentage de surface terrière coupée lors de l'intervention, pour l'ensemble des individus des trois classes obtenues lors de la CAH avec la distance de Ward sur les données de ACP sur l'ensemble des contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant le RCP 4.5.

5 Discussion

5.1 Analyse des résultats

5.1.1 Résultats de l’algorithme génétique

Le nombre de générations de l’algorithme (voir le tableau 1) est le même pour trois des jeux de données climatiques. La différence observée pour le RCP 8.5 du CNRM peut venir des processus aléatoires de l’algorithme. Pour le vérifier, il faudrait le lancer un plus grand nombre de fois. Si on regarde les scores des individus identifiés, la valeur maximale est plus grande pour les RCP 4.5 et la variance est plus faible. Le premier phénomène peut s’expliquer par le fait que ses prédictions sont moins extrêmes que les RCP 8.5. Lors d’une simulation, les contraintes sont plus facilement respectées et donc le score est plus haut. Les scores maxima obtenus avec les données du CNRM sont d’ailleurs légèrement au-dessus de ceux obtenus avec le MPI, qui est le plus extrême des deux modèles. Pour ce qui est de la variance, il est possible que les événements extrêmes, plus fréquents avec le RCP 8.5, augmentent l’influence de l’aléatoire présent dans ForCEEPS. Cependant la variance issue des simulations faites avec le modèle du CNRM est plus faible qu’avec le MPI. Il est donc plus difficile de conclure que pour le score maximum. Enfin, si le nombre d’individus viables identifiés diffère pour les quatre cas, la proportion est sensiblement la même, autour de 82%. L’algorithme ne calcule pas plus de 20% d’individu non-viable. Il parcourt donc assez efficacement l’espace des contrôles quel que soit le jeu de données climatique utilisé.

5.1.2 Résultats de l’ACP

L’ACP a permis d’identifier les liens entre les différentes variables (voir figure 7). La production est liée au nombre d’espèces. Cela peut venir du fait, qu’après une coupe, il y a plus de lumière accessible au sol ce qui favorise l’établissement d’un grand nombre d’espèces. Des coupes régulières ou importantes peuvent donc augmenter la production et le nombre d’espèces. De plus, l’étude [22] a mis en évidence qu’une grande diversité favorise la productivité d’une forêt, même si cette relation n’est pas toujours observée et dépend du contexte climatique [23]. Dans notre étude, cet effet est dû à l’augmentation de la résistance de certaines espèces arboricoles grâce à la proximité d’arbre d’une autre espèce complémentaire. Le lien entre la surface terrière et le taux de mortalité peut venir de la compétition lumineuse. Lorsque la surface terrière est grande, cela implique un grand couvert forestier car le diamètre et la taille du houppier des arbres sont liées dans le modèle. Il est donc possible que le manque de lumière entraîne la mort de plus d’arbres. La corrélation négative entre la production et la surface terrière vient du fait que les coupes diminuent logiquement le nombre d’arbres présents sur la placette et donc la surface terrière. Celle entre la production et le taux de mortalité peut s’expliquer par la diminution de la compétition entre les arbres après une coupe.

5.1.3 Analyse des itinéraires de gestion

Le premier objectif de ce stage était d’identifier des itinéraires de gestion viables dans le temps en contexte de changement climatique. En mettant en relation les figures 9, 10 et 11 il est possible de tirer des conclusions sur les itinéraires identifiés. Nous parlons ici des résultats obtenus avec les données climatiques les plus restrictives que nous avons utilisées, soit celle

du modèle MPI suivant le scénario RCP 8.5. Nous avons identifié trois classes d’itinéraire de gestion.

La première, en jaune sur les graphiques, favorise la capitalisation de la forêt par augmentation de la surface terrière. Ce sont donc des forêts dont la production n’est pas la priorité et qui présente moins d’espèces arboricoles. L’itinéraire de gestion permettant cela est basé sur des coupes plus légères que celle des autres, i.e. 25% de la surface terrière présente et espacées dans le temps tous les 25 ans. Notons qu’usuellement des périodes supérieures à 10 ans sont pratiquées sur le terrain. Ensuite, les itinéraires issues de la classe trois, en rouge, sont plus productifs et permettent une plus grande diversité d’espèce arboricole. Les itinéraires de gestion associée coupent une plus grande proportion des arbres lors des interventions, environs 35%. Pour finir, la dernière classe, en vert, correspond à des futaies irrégulière, c’est-à-dire des forêts avec une grande diversité de diamètre. La gestion y est intermédiaire sur les autres points.

Cependant, il est difficile de conclure en se concentrant sur les classes car les différences ne sont pas catégoriques. Il est possible que l’ensemble des individus identifiés par l’algorithme génétique ne soit pas assez diversifié pour construire des classes éloignées les unes des autres.

5.2 Discussion méthodologique

L’objectif de ce stage était d’explorer un ensemble de méthodes issues de la théorie du contrôle [6], puis d’en sélectionner une appropriée à notre modèle forestier pour étudier le maintien d’un système forestier riche en espèces dans des contraintes permet de minimiser les choix de gestion à faire avant les simulations.

Après une étude de la littérature [14, 15, 16], couplé à une analyse précise de la structure mathématique du modèle individu-centré forestier FORCEEPs, nous avons décidé de formaliser et de résoudre notre problème de contrôle en utilisant les algorithmes génétiques. De cette manière, contrairement aux approches classiques classiques [24], il devint possible de tester sans biais un très grand ensemble de contrôle possibles afin d’inférer a posteriori le sous-ensemble de solutions de gestion appropriées. De plus, l’utilisation de cette méthode nous a permis, non pas de trouver un contrôle optimal, mais d’identifier un grand nombre de contrôles viables. Cette manière algorithmique de coupler un modèle individu-centré avec les algorithmes génétiques, nous permet donc d’envisager pour la première fois d’appliquer l’esprit de la théorie de la viabilité de [25, 26], issue de la théorie du contrôle à des cas difficilement transformables en des problèmes purement analytiques.

Le temps qui m’était imparti, ne m’a pas permis de produire l’ensemble des résultats qui m’auraient permis de conclure définitivement, quant à quel type de contrôle permet d’optimiser quel objectifs forestier. De plus, certains résultats montrent un biais dans notre parcours de l’espace de contrôle. Par exemple, la variable de contrôle correspondant au type de coupe est le plus souvent égale à 0.5 et cela est peut être encouragé par la mutation de proche en proche. Une correction de certains biais que nous avons identifiés dans l’algorithme permettra de mieux parcourir l’espace et, peut être, de tirer des conclusions plus poussées.

Conclusion

Lors de ce stage, je me suis intéressée à la viabilité de la gestion forestière dans l’optique de développer un outil d’aide à la décision sylvicole.

La solution que nous avons retenue est l’utilisation des méthodes issues de la théorie du contrôle combiné à un modèle de simulation forestier. Notre choix s’est porté sur le modèle mécaniste individu-centré ForCEEPS (Forest Community Ecology and Ecosystem Processes). C’est un modèle de type gap-modèle adapté à la simulation sur long terme de l’évolution des forêts en prenant en compte le changement climatique. De plus, il est paramétré pour les espèces arboricoles présentes en France ce qui correspond à notre cas d’application. Pour finir, il est adapté à la gestion sylvicole car elle y est implémentée.

J’ai commencé par analyser ce modèle en détails grâce aux données de la littérature et aux codes de la plateforme Capsis. Cette première étape était essentielle pour la définition du problème de contrôle car cela m’a permis d’identifier les variables d’état du système ainsi que sa dynamique globale. J’ai ensuite pu, à partir de l’implémentation existante dans Capsis et des discussions avec une écologue, définir mathématiquement le contrôle et son influence sur la dynamique du système. Nous nous sommes ici intéressées à l’influence de la gestion forestière restreinte aux interventions de coupe. Pour finir, il nous a fallu choisir les critères de satisfaction de notre système qui m’ont permis de définir l’espace des contraintes. Cette fois encore, les choix ont été faits, en accord avec ma tutrice écologue (Marion Jourdan), pour correspondre à notre cas d’étude.

Notre première intuition était que la complexité d’un modèle comme ForCEEPS serait compensée par un espace de contrôle réduit. Cela aurait permis d’utiliser des méthodes classiques de la théorie de viabilité parcourant l’entière de l’espace de contrôle. Cependant, il se trouve que la grande combinatoire de l’espace de contrôle empêche la recherche exhaustive des contrôles satisfaisants. Nous avons donc décidé d’aborder notre problème comme un problème d’optimisation en grande dimension, non pas pour trouver un contrôle optimal, mais pour identifier un grand nombre de contrôles viables. Pour le résoudre, nous avons choisi d’utiliser un algorithme génétique. Après avoir converti notre problème de contrôle en problème génétique, j’ai implémenté un algorithme. L’étude de sensibilité des paramètres a eu pour but de trouver les paramètres optimisant le nombre et la proportion de contrôle viables identifiés. Ces deux critères permettent de quantifier l’étendue de l’espace de contrôle parcouru par l’algorithme ainsi que l’efficacité de ce parcours.

La mise en application s’est faite sur la forêt de Brun située dans l’ouest de la France. Nous avons étudié les scénarios de prédiction du rapport du GIEC de 2014 à plus 4.5 et 8.5 $W m^{-2}$, RCP 4.5 et 8.5. Les données climatiques utilisées viennent des modèles du MPI, considéré dans la moyenne, et du CNRM, moins extrême.

Pour finir, nous classons les contrôles viables identifiés avec une classification à ascendance hiérarchique basée sur les données d’une analyse en composantes principales sur les variables qui définissent les contraintes. Cela permet de regrouper les contrôles en fonction des contraintes qu’ils favorisent. Pour identifier les types de contrôle associés à chaque classe nous avons comparé une analyse qualitative des variables de contrôle. Pour les données climatiques du modèle MPI suivant de RCP 8.5 trois types d’itinéraire de gestion ont été identifiés. Une favorisera les forêts très capitalisées, une autre des forêts permettant de produire du bois avec une diversité spécifique plus forte et la dernière des forêts avec des arbres de diamètre irrégulier.

Références

- [1] Xavier MORIN et al. « Beyond forest succession : A gap model to study ecosystem functioning and tree community composition under climate change ». In : *Functional Ecology* 35 (2021), p. 955-975. DOI : [10.1111/1365-2435.13760](https://doi.org/10.1111/1365-2435.13760).
- [2] Marion JOURDAN et al. « Managing mixed stands can mitigate severe climate change impacts on French alpine forests ». In : *Regional Environmental Change* 21.78 (2021). DOI : [10.1007/s10113-021-01805-y](https://doi.org/10.1007/s10113-021-01805-y).
- [3] Harald BUGMANN. « One the ecology of mountainous forests in a changing climate : A simulation study ». Thèse de doct. ETH Zürich, 1994, p. 44-93. DOI : [10.3929/ethz-a-000946508](https://doi.org/10.3929/ethz-a-000946508).
- [4] Fischlin ANDREAS, Bugmann HARALD et Gyalistras DIMITRIOS. « Sensitivity of a forest ecosystem model to climate parametrization schemes ». In : *Environmental Pollution* 87.3 (1995), p. 267-282. DOI : [10.1016/0269-7491\(94\)P4158-K](https://doi.org/10.1016/0269-7491(94)P4158-K).
- [5] Samuel DUFOUR-KOWALSKI et al. « Capsis : an open software framework and community for forest growth modelling ». In : *Annals of Forest Science* 69 (mars 2012), p. 211-233. DOI : [10.1007/s13595-011-0140-9](https://doi.org/10.1007/s13595-011-0140-9).
- [6] Vernon L SMITH. « Control theory applied to natural and environmental resources an exposition ». In : *Journal of Environmental Economics and Management* 4.1 (1977), p. 1-24. DOI : [10.1016/0095-0696\(77\)90012-2](https://doi.org/10.1016/0095-0696(77)90012-2).
- [7] Claire BERNARD. « La théorie de la viabilité au service de la modélisation mathématique du développement durable. Application au cas de la forêt humide de Madagascar ». Thèse de doct. Université Blaise Pascal - Clermont-Ferrand II, 2011.
- [8] Charles ROUGÉ. « Resilience and vulnerability in the framework of viability theory and stochastic controlled dynamical systems ». Thèse de doct. Université Blaise Pascal - Clermont-Ferrand II, 2013.
- [9] Charles ROUGÉ, Jean-Denis MATHIAS et Guillaume DEFFUANT. « Relevance of control theory to design and maintenance problems in time-variant reliability : The case of stochastic viability ». In : *Reliability Engineering & System Safety* 132 (2014), p. 250-260. DOI : [10.1016/j.ress.2014.07.025](https://doi.org/10.1016/j.ress.2014.07.025).
- [10] R.K. PACHAURI, L. MAYER et Intergovernmental Panel on CLIMATE CHANGE (EDS.) « Climate Change 2014 : Synthesis Report ». In : *Intergovernmental Panel on Climate Change, Geneva, Switzerland* (2015).
- [11] Charles W. THORNTHWAITE et J. R. MATHER. « Instructions and tables for computing potential evapotranspiration and the water balance ». In : *Publications on Climatology* 10.3 (1957), p. 185-310.
- [12] A. WEHRLI et al. « Improving the establishment submodel of a forest patch model to assess the long-term protective effect of mountain forests ». In : *European Journal of Forest Research* 126 (jan. 2007), p. 131-145. DOI : [10.1007/s10342-006-0142-6](https://doi.org/10.1007/s10342-006-0142-6).
- [13] IGN-OCRE. URL : <https://ocre-gp.ign.fr/ocre> (visité le 06/2023).
- [14] Peter J. FLEMING et Carlos M. FONSECA. « Genetic algorithms in control systems engineering ». In : *IFAC Proceedings Volumes* 26.2 (1993), p. 605-612. DOI : [10.1016/S1474-6670\(17\)49015-X](https://doi.org/10.1016/S1474-6670(17)49015-X).

- [15] Mathias JEAN-DENIS, Balandraud XAVIER et Grediac MICHEL. « Applying a genetic algorithm to the optimization of composite patches ». In : *Computers & Structures* 84.12 (2006), p. 823-834. DOI : [10.1016/j.compstruc.2005.12.004](https://doi.org/10.1016/j.compstruc.2005.12.004).
- [16] Rafal DREZEWSKI et Krzysztof DOROZ. « An Agent-Based Co-Evolutionary Multi-Objective Algorithm for Portfolio Optimization ». In : *Symmetry* 9.168 (2017). DOI : [10.3390/sym9090168](https://doi.org/10.3390/sym9090168).
- [17] Hotelling HAROLD. « Analysis of a complex of statistical variables into principal components ». In : *Journal of Educational Psychology* 24 (1933), p. 417-441. DOI : <http://dx.doi.org/10.1037/h0071325>.
- [18] Joe H. Ward JR. « Hierarchical Grouping to Optimize an Objective Function ». In : *Journal of the American Statistical Association* 58.301 (1963), p. 236-244. DOI : [10.1080/01621459.1963.10500845](https://doi.org/10.1080/01621459.1963.10500845).
- [19] Sven KOTLARSKI et al. « Regional climate modeling on European scales : A joint standard evaluation of the EURO-CORDEX RCM ensemble ». In : *Geoscientific Model Development* 7 (juill. 2014), p. 1297-1333. DOI : [10.5194/gmd-7-1297-2014](https://doi.org/10.5194/gmd-7-1297-2014).
- [20] Carol MCSWEENEY et al. « Selecting CMIP5 GCMs for downscaling over multiple regions ». In : *Climate Dynamics* 44 (juin 2014), p. 3237-3260. DOI : [10.1007/s00382-014-2418-8](https://doi.org/10.1007/s00382-014-2418-8).
- [21] Marion JOURDAN et al. « Reliable predictions of forest ecosystem functioning require flawless climate forcings ». In : *Agricultural and Forest Meteorology* 311 (2021), p. 108703. ISSN : 0168-1923. DOI : [10.1016/j.agrformet.2021.108703](https://doi.org/10.1016/j.agrformet.2021.108703).
- [22] Xavier MORIN et al. « Tree species richness promotes productivity in temperate forests through strong complementarity between species ». In : *Ecology Letters* 14.12 (2011), p. 1211-1219. DOI : [10.1111/j.1461-0248.2011.01691.x](https://doi.org/10.1111/j.1461-0248.2011.01691.x).
- [23] Jean-Baptiste PICHANCOURT et al. « Growing biodiverse carbon-rich forests ». In : *Global change biology* 20.2 (2014), p. 382-393.
- [24] Timothy THRIPPLETON et al. « Balancing disturbance risk and ecosystem service provisioning in Swiss mountain forests : an increasing challenge under climate change ». In : *Regional Environmental Change* 23.29 (2023). DOI : [10.1007/s10113-022-02015-w](https://doi.org/10.1007/s10113-022-02015-w).
- [25] Jean-Pierre AUBIN. « A Survey of Viability Theory ». In : *SIAM Journal on Control and Optimization* 28.4 (1990), p. 749-788. DOI : [10.1137/0328044](https://doi.org/10.1137/0328044).
- [26] Jean-Pierre AUBIN, Alexandre M BAYEN et Patrick SAINT-PIERRE. *Viability theory : new directions*. Springer Science & Business Media, 2011. URL : <https://link.springer.com/book/10.1007/978-3-642-16684-6>.
- [27] Markus DIDION et al. « Gaining local accuracy while not losing generality - Extending the range of gap model applications ». In : *Canadian Journal of Forest Research* 39 (juin 2009), p. 1092-1107. DOI : [10.1139/X09-041](https://doi.org/10.1139/X09-041).

A Équations de ForCEEPS

Les équations sont tirées des articles [3], [4], et [1] et des codes de la plateforme Capsis.

A.1 Modèle environnement

A.1.1 Bilan hydrique

L'un des besoins les plus importants des plantes est l'eau. Pour calculer comment le climat influence l'eau disponible pour les arbres, il est nécessaire d'étudier les précipitations ainsi que les températures.

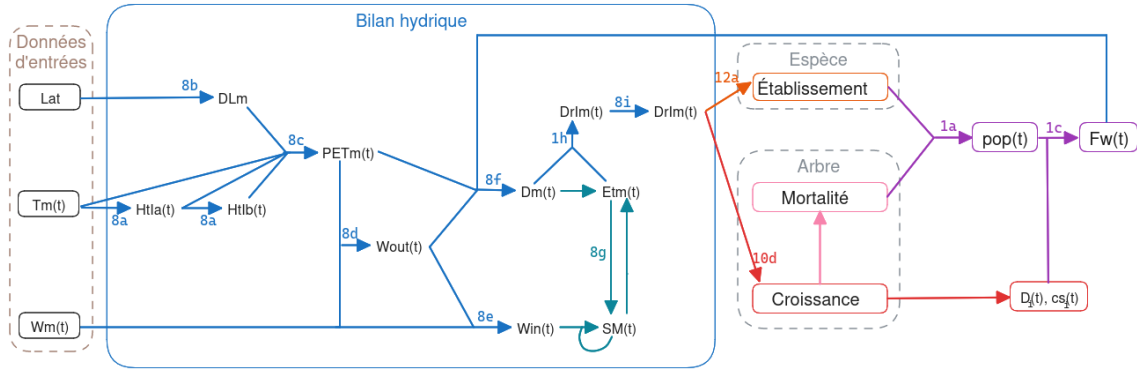


FIGURE 12 – Schéma d'organisation du bilan hydrique dans le modèle ForCEEPS, les flèches représentent les équations, les couleurs représentent les différents processus et les points beiges représentent les données d'entrée du modèle.

La première étape dans le modèle est de calculer les indices de chaleur $HtI_A(t)$ et $HtI_B(t)$ en fonction de toutes les températures $T_m(t)$ de l'année et des constantes $(c_i)_{i=1,\dots,6}$.

$$\begin{aligned} HtI_A(t) &= \sum_{m=1}^{12} \max(0, c_1 T_m(t))^{c_2} \\ HtI_B(t) &= c_3 HtI_A(t)^3 + c_4 HtI_A(t)^2 + c_5 HtI_A(t) + c_6 \end{aligned} \quad (8a)$$

Il est nécessaire de prendre en compte la longueur des journées DL_m durant chaque mois pour évaluer l'évaporation. Cela dépend de la latitude Lat et des constantes mensuelles c_{A_m} et c_{B_m} . Ce sont des paramètres qui restent constants dans le temps.

$$DL_m = c_{A_m} + c_{B_m} Lat \quad (8b)$$

L'évapotranspiration potentielle mensuelle $PET_m(t)$ est évaluée grâce à ces durées des jours DL_m , aux indices de chaleur $HtI_A(t)$ et $HtI_B(t)$, aux températures mensuelles $T_m(t)$ ainsi qu'une constante $c_{pm} = 1.6$.

$$PET_m(t) = c_{pm} \times DL_m \times \left(\frac{10}{HtI_A(t)} \max(0, T_m(t)) \right)^{HtI_B(t)} \quad (8c)$$

On s'intéresse à la quantité d'eau interceptée $W_{out_m}(t)$ par le couvert végétal durant le mois m . Elle est définie comme le minimum entre l'évapotranspiration $PET_m(t)$ et les précipitations

du mois $W_m(t)$ pondérées par la constante c_{ip} représentant la proportion moyenne d'eau arrêtée avant d'arriver au sol.

$$W_{out_m}(t) = \min(c_{ip} \times W_m(t), PET_m(t)) \quad (8d)$$

On en déduit la quantité d'eau infiltrée dans le sol $W_{in_m}(t)$.

$$W_{in_m}(t) = W_m(t) - W_{out_m}(t) \quad (8e)$$

Enfin, on peut calculer l'évaporation du sol $D_m(t)$ en utilisant la quantité totale de feuillage sur la placette $Fw(t)$ (1c).

$$D_m(t) = \left(1 + \frac{\log(Fw(t) + 1)}{20}\right) \times PET_m(t) - W_{out_m}(t) \quad (8f)$$

Le calcul de l'évapotranspiration des arbres $ET_m(t)$ se fait itérativement sur les mois de toute la simulation. Elle dépend de l'humidité du sol $SM_m(t)$ durant le mois qui est initialement égale à la surface de la réserve utile $Su = 20 \text{ m}^2$,

$$\begin{aligned} SM_1(1) &= Su \\ ET_m(t) &= \min(D_m(t), 12 \times SM_m(t)/Su) \\ SM_{m+1}(t) &= \max(\min(SM_m(t) + W_{in_m}(t) - ET_m(t), Su), 0) \end{aligned} \quad (8g)$$

Il faut noter que $SM_{12+1}(t) = SM_1(t + 1)$, cela correspond au changement d'année entre décembre ($m = 12$) et janvier ($m = 1$).

L'évaporation du sol $D_m(t)$ et celle des arbres $ET_m(t)$ permettent le calcul du stress hydrique mensuel $DrI_m(t)$.

$$DrI_m(t) = \begin{cases} 1 - \frac{ET_m(t)}{D_m(t)} & \text{si } D_m(t) \neq 0 \text{ et } T_m(t) \geq T_0 = 5.5^\circ\text{C} \\ 0 & \text{sinon} \end{cases} \quad (8h)$$

L'indice de stress hydrique annuel $DrI(t)$ est simplement obtenu grâce aux stress hydriques mensuels $DrI_m(t)$. On fait la moyenne sur les mois M_{eff_s} où la transpiration et la demande sont significatives. C'est-à-dire ceux où l'arbre a ses feuilles : toute l'année pour les feuillages persistants et d'avril à octobre pour les caduques. Attention les mois où la condition $T_m(t) \geq T_0 = 5.5^\circ\text{C}$ n'est pas respectée ne sont pas compris dans la moyenne.

$$DrI(t) = \text{moyenne}_{m \in M_{eff_s}} DrI_m(t) \quad (8i)$$

A.1.2 Variables climatiques

En plus du stress hydrique deux autres variables seront utilisées pour modéliser l'effet du climat sur les arbres. Ce sont des variables annuelles calculées grâce aux données climatiques.

A.1.2.1 Température hivernale

La température hivernale $T_W(t)$ est calculée pour l'année t à partir des températures des mois de décembre de l'année $t - 1$, de janvier et de février de l'année t .

$$T_W(t) = \min(T_{12}(t - 1), T_1(t), T_2(t)) \quad (8j)$$

A.1.2.2 Degrés jour

Le degrés jour $DD(t)$ est la somme de la température de chaque jour de l'année. On utilise la température de référence $T_0 = 5.5^\circ\text{C}$ la température moyenne de chaque mois $T_m(t)$ et le nombre de jours moyen dans un mois $N_{\text{days}} = 30.5$ jours. Pour corriger les erreurs dues à l'approximation faite avec la température moyenne, on utilise la variable $N_{\text{corr}}(t)$.

$$DD(t) = \sum_{m \in y(t)} \max(T_m(t) - T_0, 0) \times N_{\text{days}} \times N_{\text{corr}}(t) \quad (8k)$$

où

$$N_{\text{corr}}(t) = \begin{cases} 8.52 \times 10^{0.165 \times T_m(t)} & \text{si } T_m(t) \leq 5.5 \\ 187.2 \times 10^{0.0908 \times T_m(t)} & \text{si } 5.5 < T_m(t) \leq 15.5 \\ -31.8 + 2.377 \times T_m(t) & \text{si } T_m(t) > 15.5 \end{cases} \quad (8l)$$

A.2 Modèle des plantes

A.2.1 Modèle du houppier

La lumière a un rôle important dans la vie des arbres notamment via le mécanisme de photosynthèse. Il est donc important d'étudier la représentation du feuillage dans le modèle ainsi que l'influence sur la luminosité qui en découle. Un des points forts du modèle ForCEEPS est qu'il est individu centré. Cela permet entre autres, de prendre en compte le houppier de chaque arbre. En plus de cela, la représentation géométrique du houppier est plus réaliste que les autres modèles.

A.2.1.1 Modélisation de la forme du houppier

On sait que la taille du houppier peut être influencée par la présence d'autres arbres aux alentours. Notamment par les feuilles se trouvant au-dessus de la cime de l'arbre i .

On utilise l'indice maximal de surface foliaire d'une parcelle $LAI_{\text{max}} = 11.98$ et les caractéristiques d'espèce cs_{max_s} et cs_{min_s} , pour définir récursivement le rapport $cs_i(t)$ entre la taille de l'arbre et celui de son houppier avec l'indice de surface foliaire au-dessus de lui $LAI_{H_i}(t)$ (9g).

$$cs_i(t+1) = \max \left[cs_{\text{min}_s}, cs_{\text{max}_s} - (cs_{\text{max}_s} - cs_{\text{min}_s}) \left(\frac{LAI_{H_i}(t)}{LAI_{\text{max}}} \right)^2 \right] \quad (9a)$$

S'il n'y a pas beaucoup de feuille au-dessus de l'arbre $cs_i(t)$ sera proche de cs_{max_s} qui sont le rapport entre la hauteur et le houppier d'un arbre se trouvant en pleine lumière. Inversement quand $LAI_{H_i}(t)$ augmente on se rapproche de cs_{min_s} et des propriétés du houppier à l'ombre.

Ainsi pour chaque individu i on peut calculer la hauteur de houppier $cl_i(t)$ en fonction de sa taille $H_i(t)$ (10b) comme suit :

$$cl_i(t) = cs_i(t) \times H_i(t) \quad (9b)$$

On s'intéresse maintenant à la forme du houppier. Dans ForCEEPS elle est modélisée comme un cône pointé est vers le bas. L'espace vertical est discrétisé en couche, ou strate, de $\Delta x = 1$ m. Pour faciliter les calculs la hauteur du houppier $cl_i(t)$ de chaque arbre est modifiée pour qu'elle remplisse un nombre $Nbx_i(t)$ entier de couche comme sur la figure 2.

Des coefficients $r_{i,x}(t)$ représentant la proportion du feuillage dans chaque couche sont calculés comme la surface des parallélépipèdes formés par l'intersection du triangle et des couches. Pour la strate x de l'arbre i le coefficient correspondant est :

$$r_{i,x}(t) = \frac{2(Nbx_i(t) - x) - 1}{Nbx_i(t)^2} \times 1 \text{ m}^2 \quad (9c)$$

A.2.1.2 Accès à la lumière

On s'intéresse à la surface totale des feuilles $LA_i(t)$ d'un arbre. Elle dépend de son diamètre $D_i(t)$ et de son type de feuillage auquel sont associés les paramètres f'_s et a_s .

$$LA_i(t) = f'_s \times cs_i(t) \times D_i(t)^{a_s} \quad (9d)$$

On peut en déduire sa surface foliaire appartenant à la couche x $LA_{i,x}(t)$. Dans le modèle, on considère qu'il y a la même quantité de feuilles dans chaque strate. Elle dépend donc de sa surface foliaire $LA_i(t)$ de l'arbre et du nombre de couches qu'il traverse $Nbx_i(t)$.

$$LA_{i,x}(t) = \frac{LA_i(t)}{Nbx_i(t)} \quad (9e)$$

Cela nous permet de définir l'indice de surface foliaire au-dessus d'une strate $LAI_x(t)$ en considérant les feuilles de tous les arbres dont la taille $H_i(t)$ dépasse cette strate et la surface de la placette Sp .

$$LAI_x(t) = \frac{1}{Sp} \times \sum_{\substack{i \in \text{pop}(t) \\ H_i(t) \geq x + \Delta x}} LA_{i,x}(t) \quad (9f)$$

Si on prend x_i la strate contenant la cime de l'arbre, on obtient l'indice de surface foliaire au-dessus d'un arbre $LAI_{H_i}(t)$.

$$LAI_{H_i}(t) = LAI_{x_i}(t) \text{ avec } x_i < H_i(t) \leq x_i + \Delta x \quad (9g)$$

On peut aussi calculer l'indice de surface foliaire totale $LAI_{\text{tot}}(t)$ en utilisant les surfaces foliaires $LA_i(t)$ de tous les arbres et la surface de la placette Sp .

$$LAI_{\text{tot}}(t) = \frac{1}{Sp} \times \sum_{\forall i \in \text{pop}(t)} LA_i(t) \quad (9h)$$

Finalement, on obtient la lumière disponible dans chaque couche $Lav_x(t)$ et au sol $Lav(t)$ grâce au indice correspondant et à la constante c_{light} .

$$Lav_x(t) = \exp(-c_{\text{light}} LAI_x(t)) \quad (9i)$$

$$Lav(t) = \exp(-c_{\text{light}} LAI_{\text{tot}}(t)) \quad (9j)$$

A.2.2 Modèle de croissance

Dans cette partie, les diamètres sont en cm et les taille en m les conversions ne sont pas explicités mais elles sont prises en compte dans le code. Pour chaque arbre ayant survécu, on calcule sa croissance via l'augmentation de son diamètre à hauteur de poitrine. C'est un calcul qui est fait pour chaque individu i en deux étapes, voir la figure 13.

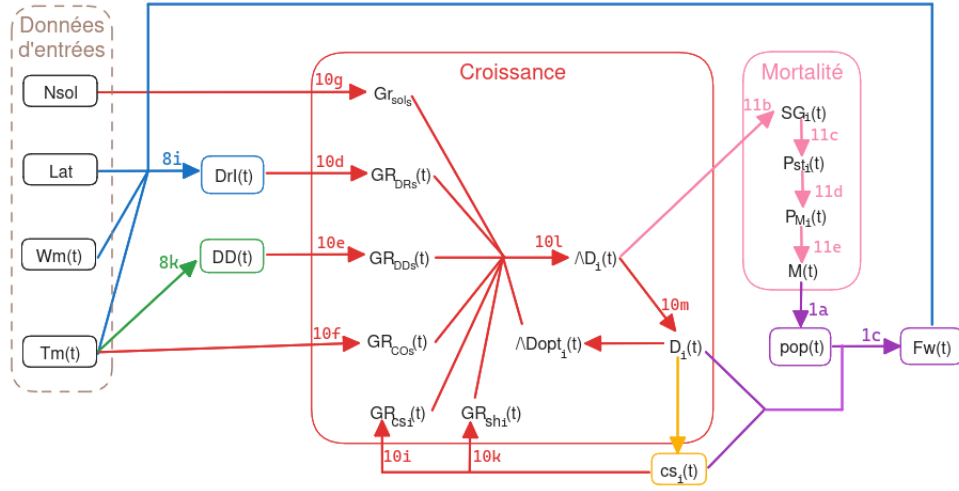


FIGURE 13 – Schéma d’organisation de la croissance et de la mortalité dans le modèle ForCEEPS, les flèches représentent les équations, les couleurs représentent les différents processus et les points beiges représentent les données d’entrée du modèle.

A.2.2.1 Croissance optimale

On peut évaluer la croissance d’un arbre en fonction des propriétés de son espèce, notamment la taille maximum totale H_{\max_s} et à partir de la poitrine b_{\max_s} .

$$b_{\max_s} = H_{\max_s} - b \quad \text{où} \quad b = 1.37 \text{ m} \quad (10a)$$

On a aussi besoin de sa taille $H_i(t)$, qu’on calcule grâce à son diamètre actuel $D_i(t)$ et à la constante spécifique s_s .

$$H_i(t) = b + b_{\max_s} \left(1 - \exp\left(-\frac{s_s}{b_{\max_s}} D_i(t)\right) \right) \quad (10b)$$

Enfin, on utilise encore une constante spécifique g_s pour estimer la croissance optimale $\Delta D_{\text{opt}_i}(t)$ de l’arbre durant l’année t .

$$\Delta D_{\text{opt}_i}(t+1) = g_s \frac{D_i(t) \left(1 - \frac{H_i(t)}{H_{\max_s}} \right)}{2H_{\max_s} - b_{\max_s} \times \exp\left(\left(\frac{-s_s}{b_{\max_s}} D_i(t)\right) \times \left(\frac{-s_s}{b_{\max_s}} D_i(t) + 2\right)\right)} \quad (10c)$$

A.2.2.2 Influence des facteurs extérieurs

Pour connaître la croissance réelle d’un arbre, il faut évaluer l’influence que peuvent avoir les différents facteurs climatiques. Le premier facteur est la sécheresse. Son impacte $GR_{Dr_s}(t)$ est calculé avec au stress hydrique $DrI(t)$ (8i) annuel et à la tolérance de l’espèce à la sécheresse $DrTol_s$.

$$GR_{Dr_s}(t) = \sqrt{\max\left(1 - \frac{DrI(t)}{DrTol_s}, 0\right)} \quad (10d)$$

L’influence de la température $GR_{DD_s}(t)$ est évaluée grâce au degré jours de l’année $DD(t)$ (8k) ainsi qu’à celui nécessaire à l’espèce DD_{\min_s} . Il fait aussi intervenir une correction $c_{DD} = \frac{1}{750^\circ\text{C}}$.

$$GR_{DD_s}(t) = \max\left(1 - \exp\left(c_{DD}(DD_{\min_s} - DD(t))\right), 0\right) \quad (10e)$$

Une seconde condition existe sur la température $GR_{CO_s}(t)$. Elle correspond à la dépendance entre la température et la capacité d'une espèce à assimiler du CO_2 . Pour la calculer, on utilise les températures maximales pour l'assimilation T_{CO_s} et pour une croissance optimale T_{opt_s} ainsi que les températures mensuelles $T_m(t)$. Comme pour le stress hydrique (8i) on ne regarde que les mois M_{eff_s} où l'arbre a son feuillage.

$$GR_{CO_s}(t) = \text{moyenne}_{m \in M_{eff_s}} \left(1 - 0.01 \exp \left(\frac{\log(99)}{T_{CO_s} - T_{opt_s}} (T_m(t) - T_{opt_s}) \right) \right) \quad (10f)$$

La quantité de nutriments dans le sol N_{soil} a aussi un impact GR_{soil_s} sur le développement d'un arbre. Il peut être caractérisé par des paramètres $N1_s$ et $N2_s$ qui dépendent de l'espèce et qui représentent le besoin de l'arbre en azote.

$$GR_{soil_s} = \max \left(1 - \exp \left(N1_s (N_{soil} - N2_s) \right), 0 \right) \quad (10g)$$

Enfin, c'est la photosynthèse qui permet aux arbres de créer de la matière organique. La lumière est donc un facteur important pour la croissance et c'est grâce au feuillage de l'arbre qu'elle est captée. Les espèces ont un point de compensation de la lumière LCP_s . Il va dépendre de leur tolérance à l'ombre $ShTol_s$, des tolérances à l'ombre maximum $ShTol_{max}$ et minimum $ShTol_{min}$ et des points de compensation respectif LCP_{max} et LCP_{min} .

$$ShTol_{max} = 1 \text{ et } ShTol_{min} = 9$$

$$LCP_{max} = 11.98 \text{ et } LCP_{min} = 10.10$$

$$LCP_s = LCP_{max} - (LCP_{max} - LCP_{min}) \times \frac{ShTol_s - ShTol_{max}}{ShTol_{min} - ShTol_{max}} \quad (10h)$$

Le rapport entre la hauteur et la taille du houppier $cs_i(t)$ (9a) et les paramètres d'espèces cs_{min_s} et cs_{max_s} nous permet de calculer $GR_{cs_i}(t)$, l'influence feuillage d'un arbre sur sa croissance.

$$GR_{cs_i}(t) = \min \left(\frac{4}{3} \times \frac{cs_i(t)}{cs_{max_s}} \times \frac{2LCP_s}{LCP_{max} - LCP_{min}}, 1 \right) \quad (10i)$$

La dernière étape est de prendre en compte la compétition entre les arbres pour la lumière. On utilise la lumière accessible $Lav_x(t)$ (9i) pour définir dans chaque couche les bornes de luminosité $L_{max_x}(t)$ et $L_{min_x}(t)$.

$$L_{max_x}(t) = 1 - \exp \left(-4.64(Lav_x(t) - 0.05) \right)$$

$$L_{min_x}(t) = 2.24 \left(1 - \exp \left(-1.136(Lav_x(t) - 0.08) \right) \right)$$

Cela permet de calculer $GR_{sh_x}(t)$, le facteur de réduction de la croissance en fonction de la lumière disponible dans chaque strate.

$$GR_{sh_x}(t) = \max \left(L_{max_x}(t) - (L_{max_x}(t) - L_{min_x}(t)) \times \frac{ShTol_s - ShTol_{max}}{ShTol_{min} - ShTol_{max}}, 0 \right) \quad (10j)$$

Pour finir, on somme sur les $Nbx_i(t)$ couche qu'occupe le houppier en utilisant les coefficients de proportion du feuillage $r_{i,x}(t)$ (9c). Ainsi, on obtient l'influence de l'ombre $GR_{sh_i}(t)$ sur la croissance de chaque arbre.

$$GR_{sh_i}(t) = \sum_{x=1}^{Nbx_i(t)} r_{i,x}(t) GR_{sh_x}(t) \quad (10k)$$

A.2.2.3 Croissance réelle

Pour connaître la croissance réelle du diamètre $\Delta D_i(t)$ de l'arbre i il faut diminuer la croissance optimale $\Delta D_{\text{opt}_i}(t)$ (10c) avec les facteurs calculés ci-dessus.

$$\Delta D_i(t+1) = \Delta D_{\text{opt}_i}(t+1) \left(GR_{Dr_s}(t) \times GR_{DD_s}(t) \times GR_{CO_s}(t) \times GR_{\text{soil}_s} \times GR_{\text{sh}_i}(t) \right)^{\frac{1}{3}} GR_{\text{cs}_i}(t+1) \quad (10l)$$

On obtient au final un nouveau diamètre de l'individu i .

$$D_i(t+1) = D_i(t) + \Delta D_i(t+1) \quad (10m)$$

A.2.3 Modèle de mortalité

La première étape de la simulation est l'identification des arbres morts durant l'année t . Deux effets rentrent en compte dans la mortalité. Le premier est lié à la mort naturelle d'un arbre dû à son âge. Elle est considérée uniforme durant toute sa vie. La probabilité de mortalité durant une année P_{age_s} est inversement proportionnelle à l'âge maximal de son espèce A_{max_s} et est pondérée par le paramètre $c_{\text{mort}} = 4.605$.

$$P_{\text{age}_s} = \frac{c_{\text{mort}}}{A_{\text{max}_s}} \quad (11a)$$

La pression subie à cause de la concurrence ou des facteurs abiotiques a aussi son influence sur la mortalité des individus. Cela est modélisé par une loi inversement proportionnelle à la croissance de l'arbre. On définit comme paramètres une probabilité p_{sg} et un seuil c_{yearSg} et on garde en mémoire le nombre d'années consécutives $SG_i(t)$ où la croissance de l'arbre $\Delta D_i(t)$ est inférieure au seuil $d_{\text{sg}} = 0.3 \text{ cm}$.

$$SG_i(t+1) = \begin{cases} SG_i(t) + 1 & \text{si } \Delta D_i \leq d_{\text{sg}} \\ 0 & \text{sinon} \end{cases} \quad (11b)$$

Ainsi on peut calculer $P_{\text{st}_i}(t)$, la probabilité de mortalité dû au stress.

$$P_{\text{st}_i}(t) = \begin{cases} p_{\text{sg}} & \text{si } SG_i(t) \geq c_{\text{yearSg}} \\ 0 & \text{sinon} \end{cases} \quad (11c)$$

Pour finir la probabilité de mort $P_{M_i}(t)$ est calculée pour chaque individus i en combinant la mort due à l'âge P_{age_s} et celle du stress $P_{\text{st}_i}(t)$.

$$P_{M_i}(t) = P_{\text{age}_s} + P_{\text{st}_i}(t) - P_{\text{age}_s} \times P_{\text{st}_i}(t) \quad (11d)$$

On définit $M(t)$ l'ensemble des individus de la population $\text{pop}(t)$ mort l'année t :

$$M(t) = \{ i \mid P_{M_i}(t) \geq k_M(t, i) \} \quad (11e)$$

ou $k_M(t, i)$ est un nombre aléatoire compris entre 0 et 1. On peut aussi définir $A(t)$ l'ensemble des survivants :

$$A(t) = \text{pop}(t) \setminus M(t) \quad (11f)$$

A.2.4 Modèle d'établissement

Pour finir, il faut calculer les nouveaux arbres qui arrivent dans la simulation, voir la figure 14. L'établissement d'un arbre se fera s'il supporte les conditions de l'année. Cela dépend donc des variables abiotiques de ses propriétés intrinsèques définies par son espèce. On le modélise par une probabilité d'établissements par espèces.

A.2.4.1 Calcul des seuils d'acceptabilité

On commence par définir différents seuils d'acceptabilité correspondant aux besoins des semis. Le premier est le seuil de sécheresse $P_{Dr_s}(t)$. Il compare l'indice de stress hydrique $DrI(t)$ (8i) annuel et la tolérance de l'espèce à la sécheresse $DrTol_s$.

$$P_{Dr_s}(t) = \begin{cases} 1 & \text{si } DrI(t) \leq DrTol_s \\ 0 & \text{sinon} \end{cases} \quad (12a)$$

Il existe deux conditions sur la température. La première est le seuil de germination $P_{TW_s}(t)$. Il dépend de si la température hivernale $T_W(t)$ (8j) de l'année en question est dans l'intervalle de température $T_{w_{\min_s}}, T_{w_{\max_s}}$ permettant la germination d'une graine de l'espèce s .

$$P_{TW_s}(t) = \begin{cases} 1 & \text{si } T_{w_{\min_s}} \leq T_W(t) \leq T_{w_{\max_s}} \\ 0 & \text{sinon} \end{cases} \quad (12b)$$

La seconde est le seuil du degré jours $P_{DD_s}(t)$. Il dépend de celui de l'année $DD(t)$ (8k) ainsi qu'à celui nécessaire à l'espèce DD_{\min_s} .

$$P_{DD_s}(t) = \begin{cases} 1 & \text{si } DD(t) \geq DD_{\min_s} \\ 0 & \text{sinon} \end{cases} \quad (12c)$$

Il est possible de prendre en compte l'impacte des animaux. Pour cela, il faut calculer l'effet de l'abrutissement sur les semis Br_{eff_s} en fonction la sensibilité de l'espèce Br_s et de l'intensité l'abrutissement sur le site BrI .

$$Br_{\text{eff}_s} = \begin{cases} BrI^4 & \text{si } Br_s = 1 \\ BrI^2 & \text{si } Br_s = 2 \\ BrI & \text{si } Br_s = 3 \\ BrI^{0.5} & \text{si } Br_s = 4 \\ BrI^{0.25} & \text{si } Br_s = 5 \end{cases} \quad (12d)$$

On obtient le seuil d'abrutissement P_{Br_s} en compare Br_{eff_s} à k_{Br} un nombre tiré aléatoirement entre 0 et 1.

$$P_{Br_s} = \begin{cases} 1 & \text{si } Br_{\text{eff}_s} \leq k_{Br} \\ 0 & \text{sinon} \end{cases} \quad (12e)$$

Pour finir, il y a le seuil de luminosité qui compare la lumière accessible $Lav(t)$ (9j) avec le besoin des semis $ShTolSeed_s$.

$$P_{Lav_s}(t) = \begin{cases} 1 & \text{si } Lav(t) \geq ShTolSeed_s \\ 0 & \text{sinon} \end{cases} \quad (12f)$$

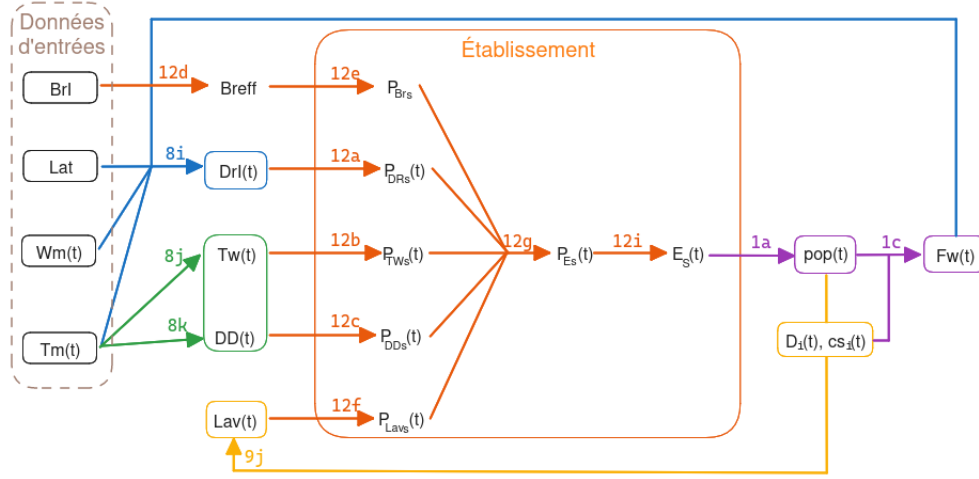


FIGURE 14 – Schéma d'organisation de l'établissement dans les modèles ForCEEPS et ForClim, les flèches représentent les équations, les couleurs représentent les différents processus et les points beiges représentent les données d'entrée du modèle.

A.2.4.2 Calcul de l'établissement

Grâce aux seuils cités ci-dessus et à un paramètre p_E on peut calculer la probabilité d'établissements $P_{E_s}(t)$ de l'espèce s durant l'année t .

$$P_{E_s}(t) = P_{D_{r_s}}(t) \times P_{TW_s}(t) \times P_{DD_s}(t) \times P_{Br_s} \times P_{Lav_s}(t) \times p_E \quad (12g)$$

Il y aura effectivement établissement d'arbre de l'espèce s durant l'année t si la probabilité est supérieure à k_E est un nombre aléatoire compris entre 0 et 1.

$$P_{E_s}(t) \geq k_E \quad (12h)$$

Pour finir, le nombre d'arbres $nb_{E_s}(t)$ de cette espèce établis l'année t est tiré selon une loi uniforme qui dépend du taux maximum d'établissement défini E_{\max} , de la taille de la placette Sp et du besoin en lumières de l'espèce $ShTol_s$.

$$nb_{E_s}(t) \sim \mathcal{U}\left([1, E_{\max} \times Sp \times ShTol_s]\right) \quad (12i)$$

Chacun des $nb_{E_s}(t)$ nouveaux individus ont réussi à s'établir. Ils intègrent la simulation avec un diamètre initial $D = D_E = 1.27$ cm et une hauteur de minimum $H = b = 1.37$ m. On note $E(t)$ l'ensemble des individus, toute espèce confondue, qui se sont établis l'année t .

B Modèle ForClim

Le modèle ForClim a été développé dans les années 1990 en se basant sur les forêts de montagnes suisses [3]. Il a peu à peu été amélioré grâce à différents travaux, notamment [12] qui propose une représentation plus détaillée de la régénération des arbres ou bien [27] qui permet d'augmenter sa précision locale. Son organisation, représentée dans la figure 15, est basée sur la division du modèle en trois modules : l'environnement, les plantes et le sol. Les flèches représentent les interactions de ces sous-modèles les uns avec les autres qui se font grâce à des variables.

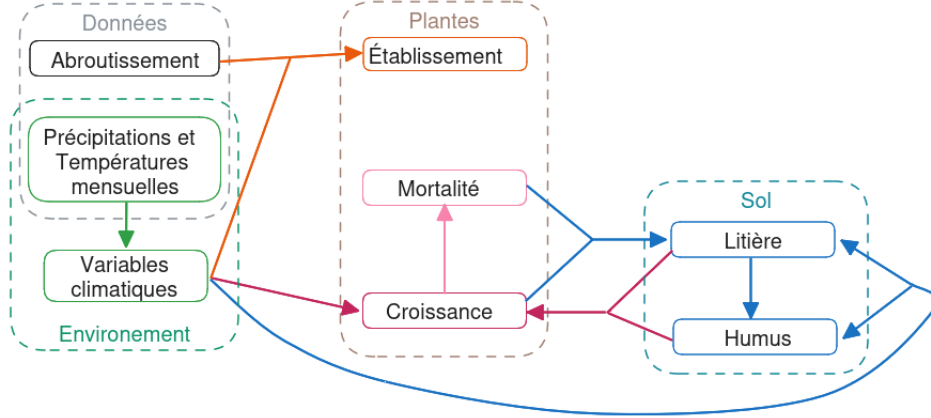


FIGURE 15 – Schéma organisation du modèle ForClim

B.1 Modèle environnement

Comme pour ForCEEPS les données climatiques $T_m(t)$ et $W_m(t)$ sont mensuelles. Elles permettent de calculer la température en hiver $T_W(t)$ et le degré-jour $DD(t)$.

B.2 Modèle des plantes

La croissance d'un arbre est déterministe. Elle dépend des données abiotiques et de son espèce. Tous les individus d'une même espèce établis la même année auront une taille identique tout au long de leur vie. Cet ensemble d'arbres identiques est appelé une cohorte. Elle est caractérisée par l'âge et l'espèce des individus qui la composent. La croissance peut alors être calculée une seule fois pour l'intégralité des arbres de la cohorte. On appelle ça la simplification par cohorte (*coghort-base*).

On s'intéresse à la croissance du diamètre à hauteur de poitrine $D(t)$. Son calcul est fait en deux étapes. D'abord, on regarde la croissance optimale $\Delta D_{\text{opt}_i}(t)$. Elle est proportionnelle à l'air du feuillage $LA(t)$ de l'arbre. Dans ForClim il est représenté comme un disque en haut de la cime. Son rayon dépend allométriquement du diamètre et il n'a pas de profondeur.

Pour obtenir la croissance réelle des contraintes abiotiques supplémentaire entre en jeu. La sécheresse avec le $DrI(t)$ et les températures avec le $DD(t)$ sont estimé dans le sous-modèle environnement. La quantité de nutriment disponible $N_{\text{soil}}(t)$ calculée dans le modèle du sol a aussi un effet. Pour finir, on prend aussi en compte la lumière disponible. Elle dépend de l'ombre créée par le feuillage des cohortes dont la hauteur dépasse celle étudiée.

On obtient ainsi $\Delta D(t)$, la croissance réelle du diamètre de la cohorte durant l'année t . On peut calculer le nouveau diamètre de la cohorte $D(t + 1) = D(t) + \Delta D(t)$.

Le modèle de la mortalité et d'établissement sont définis comme dans ForCEEPS (voir sections 1.2.2.2 et 1.2.2.3). Contrairement à la croissance, ils ne sont pas calculés pour toute la cohorte mais pour chaque arbre. Ce sont des processus partiellement stochastiques.

B.3 Modèle du sol

Le modèle du sol est basé sur le cycle de l'azote qui est le nutriment principal des plantes. La matière organique au sol est appelée litière. Son apport annuel vient des arbres morts mais aussi de branche et de feuilles qui tombent au sol. Pour calculer ces dernières, on se base sur le modèle de croissance. Quand elle se décompose l'azote qu'elle contient se minéralise et crée de l'humus chargé en azote disponible pour les plantes et leur croissance. La décomposition de la litière est influencée par le climat. En modélisant la quantité de litière et d'humus au sol, on est capable d'estimer la quantité de nutriment disponible $N_{\text{soil}}(t)$ qui permet la croissance des cohortes.

B.4 Les limites

Si ce modèle est un bon compromis entre précision et rapidité de calcul, il a plusieurs points faibles [1]. D'abord, la forme du feuillage, qu'on appelle le houppier ou le houppier, est représenté de façon simpliste : un disque dont la position en haut de l'arbre et le rayon dépendent allométriquement du diamètre. Cela implique une approximation lors du calcul de la croissance et l'établissement qui dépendent de la lumière accessible. De plus, la simplification par cohorte permet des simulations plus courtes en diminuant le nombre de calculs, mais entraîne un manque de précision. Elle empêche l'individualisation notamment dans la simulation de la concurrence pour la lumière.

C Test du nombre de classe pour la CAH

Pour tester si le nombre de classes que nous avons choisi est pertinent, nous avons comparé les résultats obtenus avec trois et quatre classes sur les résultats venant du GA avec les données climatiques venant du modèle MPI suivant le RCP 8.5.

Dans la figure 16, quand il y a quatre classes au lieu de trois, les différences entre les individus de la classe trois (en rouge) et quatre (en violet) sont plus compliquées à identifier que celle entre les autres classes. Les individus de la classe trois ont un meilleur équilibre entre la surface terrière et le taux de mortalité d'un côté, et la production et le nombre d'espèces de l'autre. Les individus de la classe quatre, quant à eux, favorisent grandement la production et le nombre d'espèces. Cependant la limitation entre les classes semble plus floue qu'avec seulement trois classes.

L'analyse des interquartiles, présentés dans la figure 17, confirme que la troisième classe est plus équilibrée que la classe quatre qui sont au-dessus pour la production et le nombre d'espèces et en dessous pour la surface terrière. De plus, la différence entre la classe trois et quatre est moins remarquable que celle entre les classes dans le cas où il n'y en a que trois. Cette quatrième classe n'ajoute pas d'information essentielle.

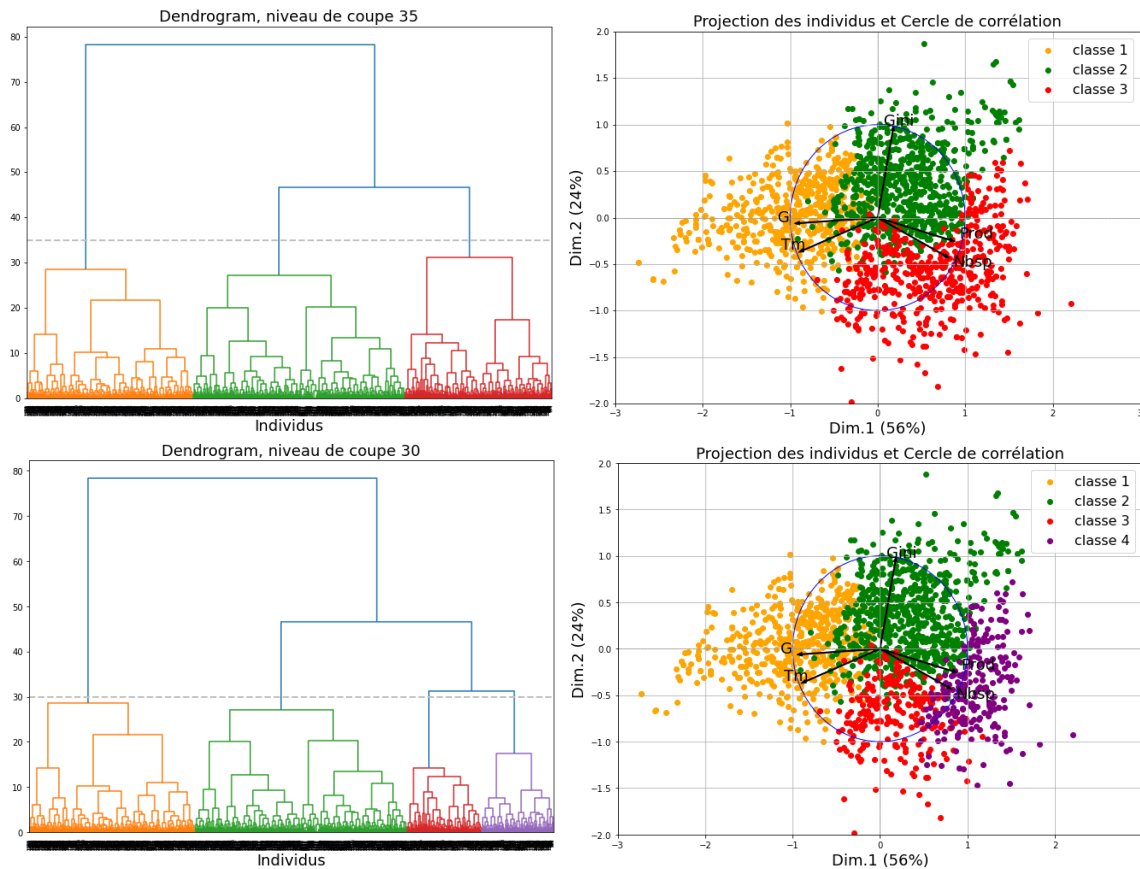


FIGURE 16 – Résultats de la CAH avec la distance de Ward sur les données de ACP sur l'ensemble des contraintes des individus viables identifiés par le GA avec les données climatiques du modèle MPI suivant le RCP 8.5. Dendrogramme (à gauche) et projection des individus et du cercle de corrélation des contraintes sur les deux composantes principales (à droite). Les couleurs correspondent aux classes d'individus, les pointillés gris sont le niveau de coupe du dendrogramme, les flèches représentent les coefficients de corrélation entre les variables et les deux composantes principales.

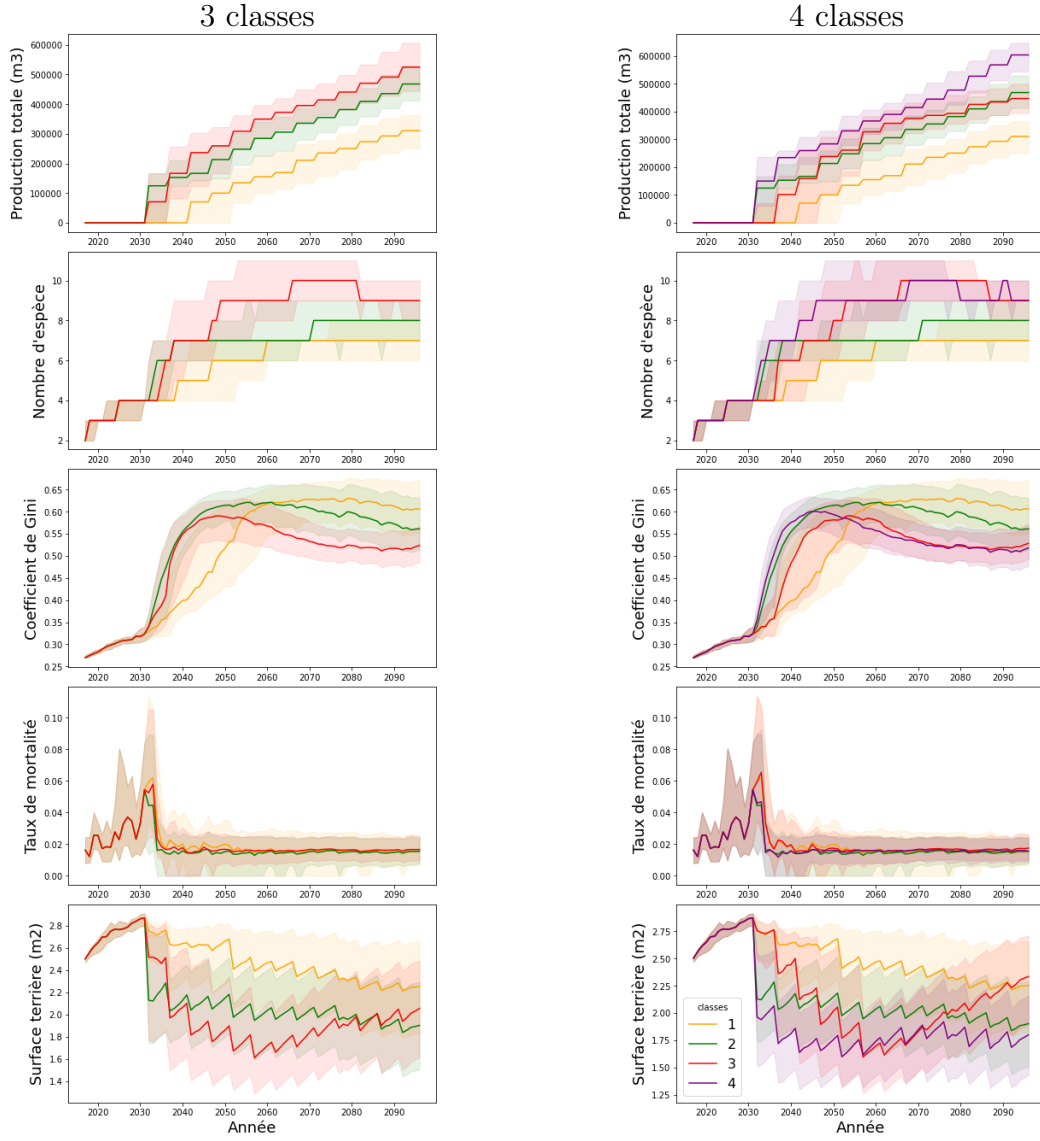


FIGURE 17 – Médiane et écart interquartile (0.25-0.75) de l'évolution au cours de la simulation des contraintes, production, nombre d'espèce, coefficient de Gini, taux de mortalité et surface terrière, pour les différentes classes obtenues lors de la CAH avec la distance de Ward sur les données de ACP sur l'ensemble des contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant le RCP 8.5.

D Figures complémentaires de l'analyse des résultats

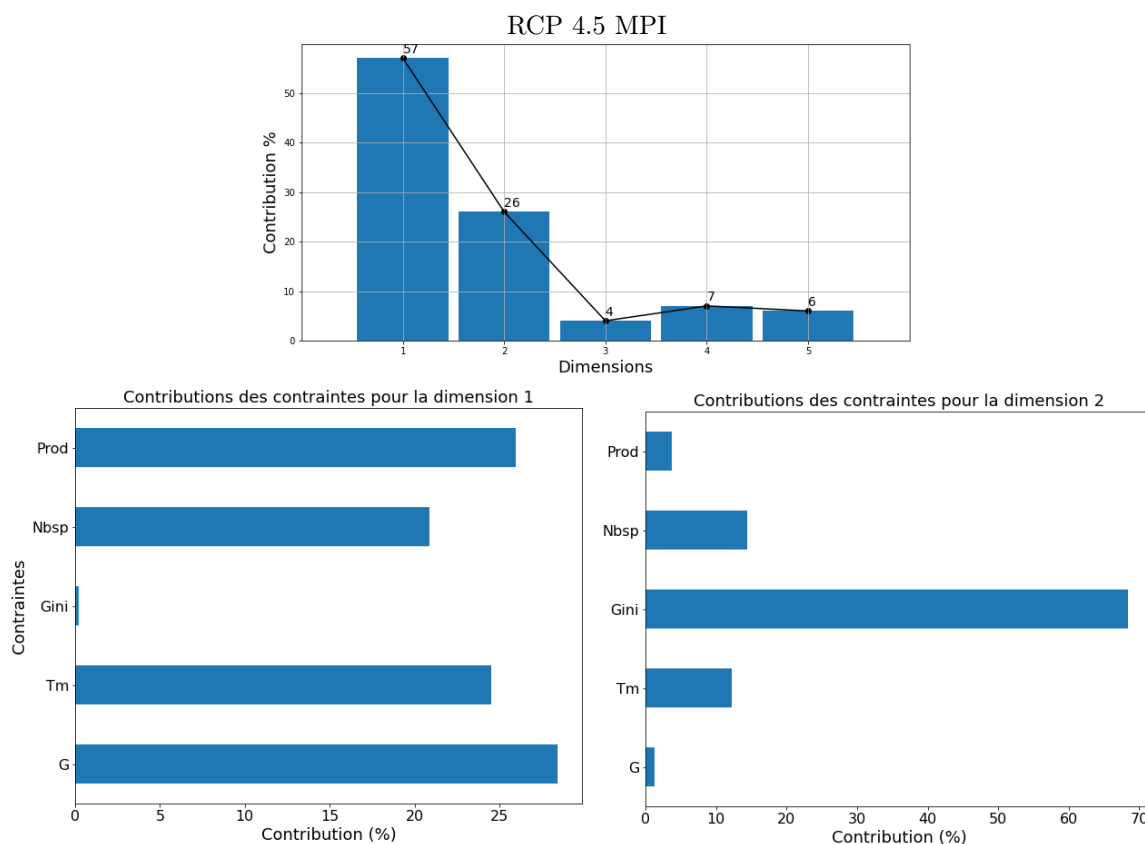


FIGURE 18 – Contribution des dimensions (en haut) calculées avec l'ACP sur les contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant le RCP 4.5. Contribution des contraintes pour les deux dimensions principales (en bas).

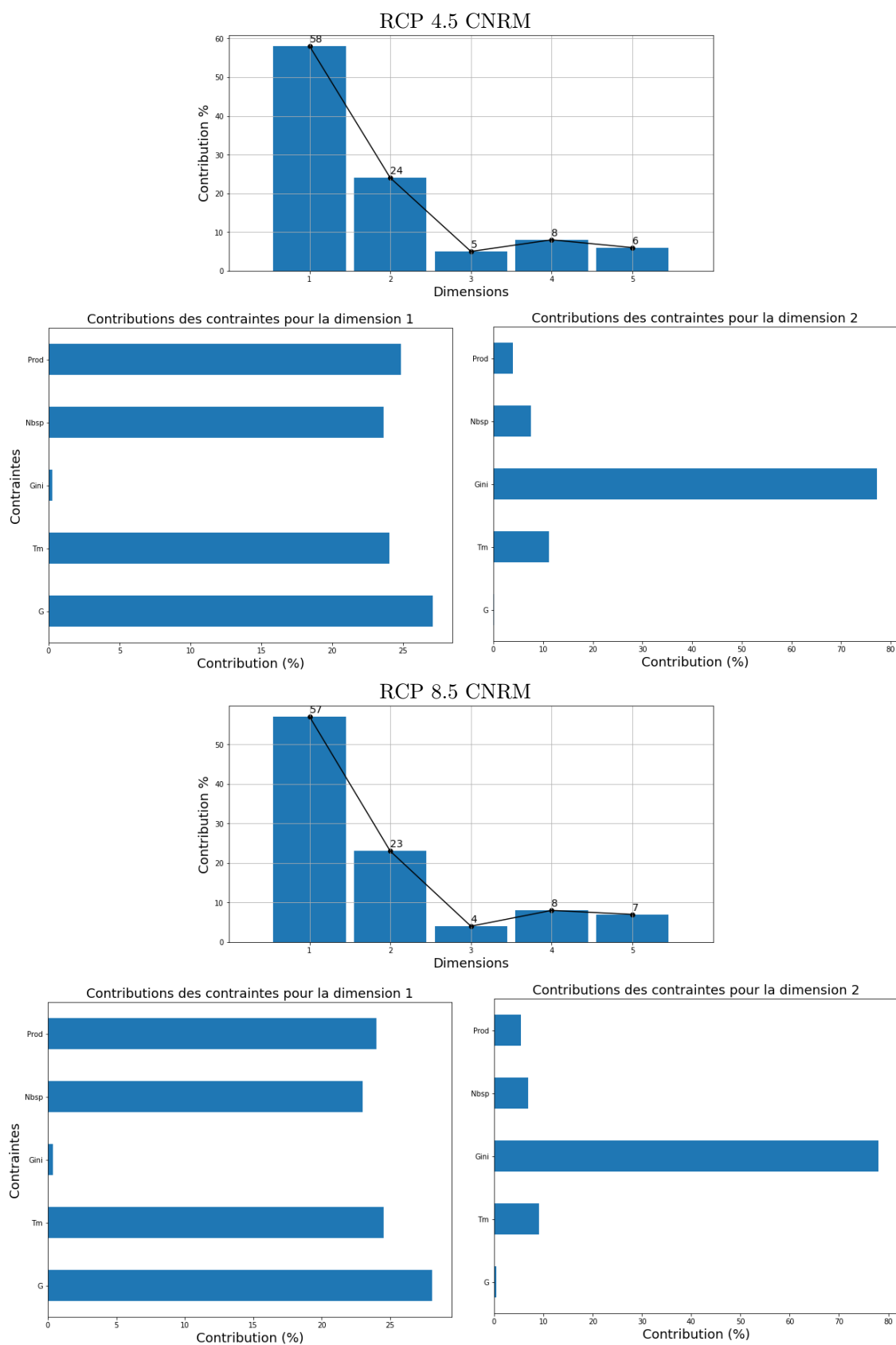


FIGURE 19 – Contribution des dimensions (en haut) calculées avec l’ACP sur les contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du CNRM suivant les RCP 4.5 et 8.5. Contribution des contraintes pour les deux dimensions principales (en bas).

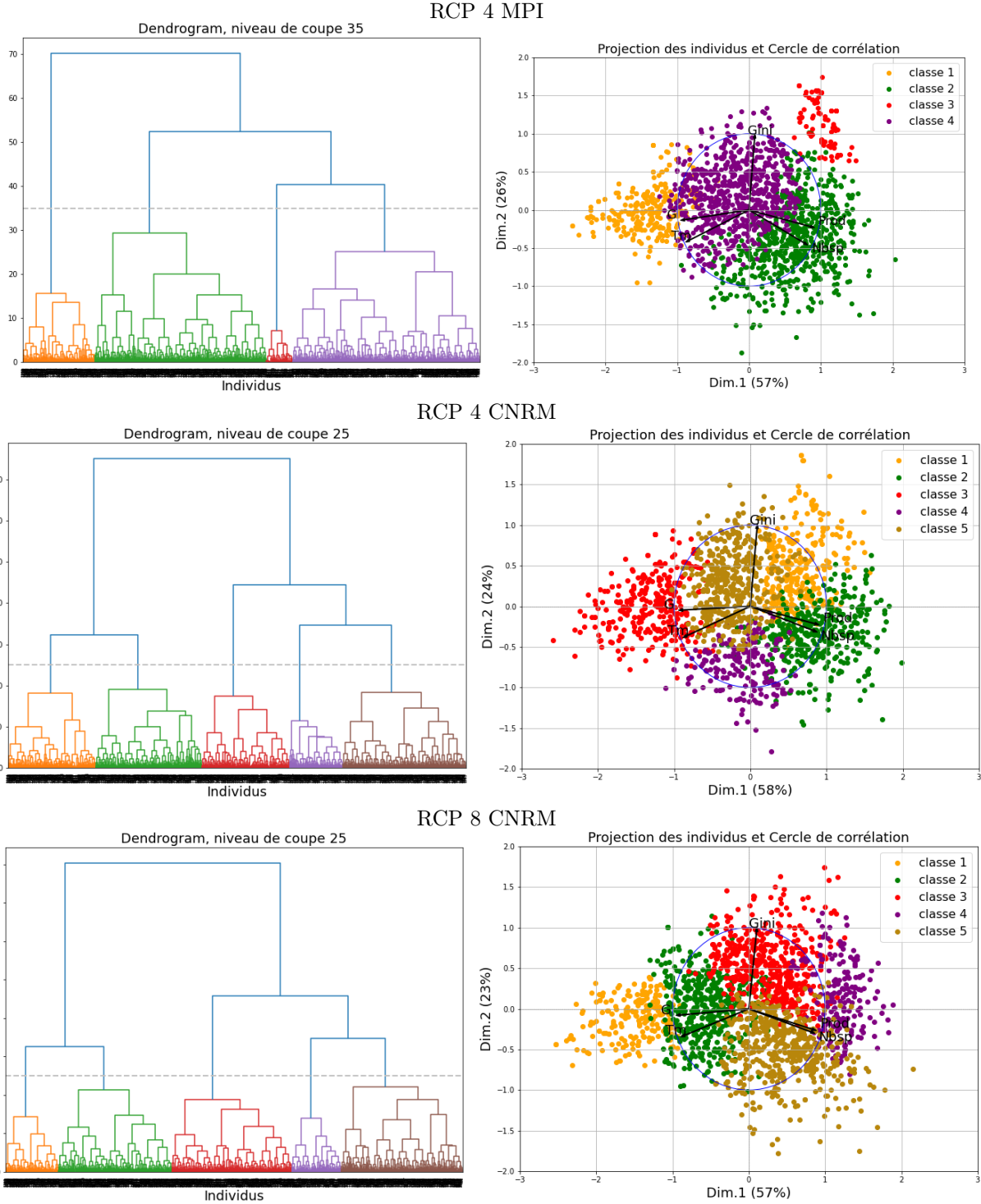


FIGURE 20 – Résultats de la CAH avec la distance de Ward sur les données de l’ACP sur l’ensemble des contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant les RCP 4.5 et du CNRM suivant les RCP 4.5 et 8.5. Le dendrogramme (à gauche) représente les individus en abscisse et leurs distances en ordonnée. La projection des individus et du cercle de corrélation des contraintes sur les deux composantes principales (à droite). Les couleurs correspondent aux classes d’individus, les pointillés gris sont le niveau de coupe du dendrogramme, les flèches représentent les coefficients de corrélation entre les variables et les deux composantes principales.

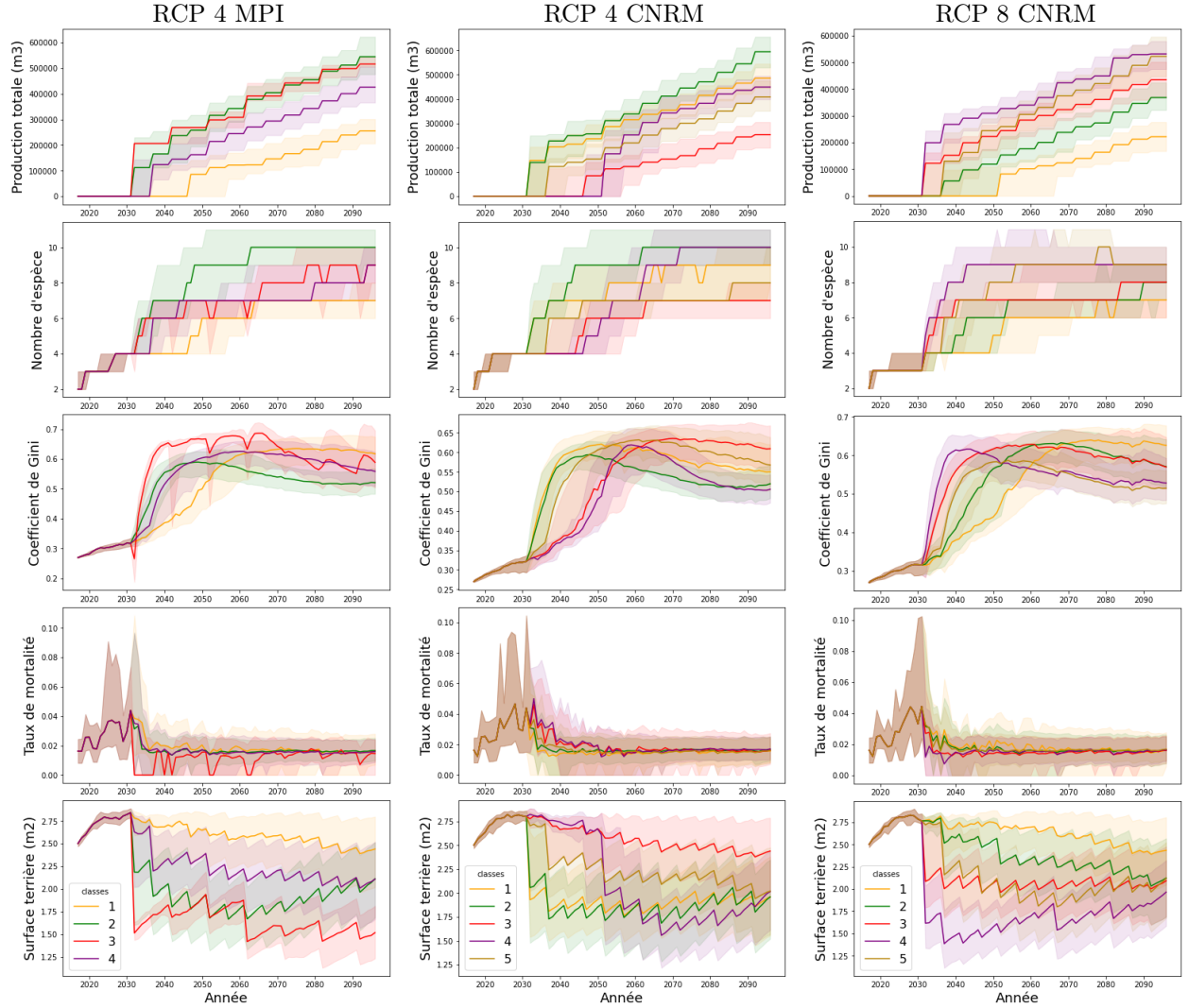


FIGURE 21 – Médiane et écart interquartile (0.25-0.75) de l'évolution au cours de la simulation des contraintes, production, nombre d'espèce, coefficient de Gini, taux de mortalité et surface terrière, pour les différentes classes obtenues lors de la CAH avec la distance de Ward sur les données de ACP sur l'ensemble des contraintes des individus viables identifiés par le GA avec les données climatiques du modèle du MPI suivant les RCP 4.5 et du CNRM suivant les RCP 4.5 et 8.5.