

VetAgro Sup

Mémoire de fin d'études d'ingénieur

Représentativité climatique du réseau
de dispositifs en jardins communs
GEN4X et particulièrement des
dispositifs méditerranéens sur le
territoire français dans un contexte de
réchauffement global

Gaspard Rihm
Option AEST
2021

VetAgro Sup

Mémoire de fin d'études d'ingénieur

Représentativité climatique du réseau
de dispositifs en jardins communs
GEN4X et particulièrement des
dispositifs méditerranéens sur le
territoire français dans un contexte de
réchauffement global

Gaspard Rihm
Option AEST
2021

Tuteur de stage : Bruno FADY, directeur de recherches à l'URFM, INRAE
Provence-Alpes-Côte d'Azur

Enseignant référent : Antoine TARDIF, VetAgro Sup

« L'étudiant conserve la qualité d'auteur ou d'inventeur au regard des dispositions du code de la propriété intellectuelle pour le contenu de son mémoire et assume l'intégralité de sa responsabilité civile, administrative et/ou pénale en cas de plagiat ou de toute autre faute administrative, civile ou pénale. Il ne saurait, en cas, seul ou avec des tiers, appeler en garantie VetAgro Sup. »

Remerciements

Je tiens en premier lieu à remercier l'URFM, qui m'a permis de réaliser ce stage, en adressant mes remerciements particuliers à Bruno Fady, pour son aide, sa disponibilité, sa bienveillance et ses conseils, qui m'ont permis non seulement de mener à bien cette mission de stage mais aussi d'en apprendre beaucoup au cours de ces 6 mois passés en unité de recherche.

Je remercie Denis Vauthier qui m'a permis d'en apprendre plus sur le cœur de mon sujet de stage.

Je souhaite également remercier les équipes de l'URFM et l'UEFM pour leur accueil et leur expertise.

J'adresse mes remerciements aux partenaires extérieurs qui m'ont beaucoup apporté pour cette mission : je remercie Marta Benito Garzón, Maurizio Marchi et Wilfried Thuiller pour leur expertise et leurs conseils, je remercie également Dominique Veisse et Frédéric Bernier qui m'ont apporté des données nécessaires à cette étude.

Enfin, je souhaite remercier Antoine Tardif pour sa disponibilité et son accompagnement pendant ce stage.

Table des matières

Introduction	1
Etat de l'art	1
1. Contexte scientifique	3
1.1. Quelques éléments d'analyse bibliométrique	3
1.1.1. Méthode	3
1.1.2. Tendances temporelles	3
1.1.3. Implication des pays.....	4
1.1.4. Écosystèmes concernés	5
1.2. Les dispositifs en jardin commun	5
1.3. Méthodologie du jardin commun forestier	6
1.4. Les Species Distribution Models, un outil puissant.....	8
2. Le réseau GEN4X.....	8
Matériel et méthodes	10
1. Aire d'étude	10
2. Partenaires et acteurs	10
3. Données de présence des dispositifs.....	10
3.1. Préparation des données.....	10
3.2. Format des données	11
4. Données environnementales	11
5. Analyse descriptive et exploratoire	14
5.1. Description de la base	14
5.2. Exploration des données climatiques.....	14
6. Modélisation des aires de compatibilité climatique du réseau GEN4X	14
6.1. Principe	14
6.2. Sélection des données environnementales	15
6.3. Modèles utilisés	16
6.4. Démarche de modélisation.....	17
7. Comparaison avec le réseau GEN4X	18

Résultats	20
1. Analyse de la base de données GEN4X	20
1.1. Description de la base	20
1.2. Analyse spatiale par sylvoécotérrégions	21
2. Exploration des données climatiques.....	22
2.1. Diagrammes température/précipitations	22
2.2. Diagrammes de saisonnalité	23
2.3. Diagrammes des valeurs extrêmes.....	24
3. Modélisation des aires de compatibilité climatique du réseau GEN4X	25
3.1. Sélection des variables d'intérêts.....	25
3.2. Résultats de BIOMOD.....	23
3.2.1. Modèles individuels	26
3.2.2. Modèles d'ensemble	27
3.2.3. Projections.....	27
3.2.4. Évolution de la distribution (Species Range Change)	28
3.3. Représentativité climatique du réseau GEN4X	30
Discussion	32
1. Performance des modèles	32
1.1. Réseau entier (jeu de données FRANCE).....	32
1.2. Dispositifs au climat méditerranéen (jeu de données GRECO)	33
1.3. Dispositifs aux espèces méditerranéennes (jeu de données ESP).....	34
2. Représentativité climatique du réseau GEN4X	34
2.1. Réseau entier.....	34
2.2. Partie méditerranéenne du réseau	35
2.3. Dispositifs à espèces méditerranéennes	35
3. Implications de ces résultats et perspectives de recherche	36
Conclusion	39

Table des figures

Figure 1 : Localisation des différents dispositifs GEN4X	7
Figure 2 : Photographies d'un dispositif en jardin commun du réseau GEN4X	7
Figure 3 : Carte du jeu de données FRANCE	11
Figure 4 : Carte du jeu de données GRECO	12
Figure 5 : Carte du jeu de données ESP	13
Figure 6 : Histogramme du nombre de dispositifs GEN4X plantés au cours du temps	20
Figure 7 : Proportion de genres représentés dans les dispositifs GEN4X	21
Figure 8 : Nombre de dispositifs expérimentaux GEN4X par sylvoécotons	21
Figure 9 : Diagrammes température/précipitation	22
Figure 10 : Diagrammes de saisonnalité	23
Figure 11 : Diagramme d'évènements extrêmes.....	24
Figure 12 : Pourcentage de variance expliqué par les 10 premières dimensions de l'ACP	25
Figure 13 : Cercle des corrélations	25
Figure 14 : Moyennes des scores ROC et TSS pour chaque algorithme BIOMOD	26
Figure 15 : Cartes des probabilités de compatibilité climatiques.....	27
Figure 16 : Cartes de « Species Range Change »	29

Table des tableaux

Tableau 1 : Scores ROC et TSS pour chaque modèle d'ensemble	27
Tableau 2 : Proportion de la surface du territoire français représentant l'aire de compatibilité climatique	28
Tableau 3 : Statistiques associées aux cartes de changement de distribution	29
Tableau 4 : Nombre de dispositifs en jardins communs du réseau GEN4X présents dans l'aire de compatibilité climatique	30
Tableau 5 : : Nombre de dispositifs en jardins communs du réseau GEN4X présents dans les 4 types de zones des cartes de « Species range change ».....	31

Liste des abréviations

ACP : Analyse en Composante Principale

CNRS : Centre National de la Recherche Scientifique

CNPF : Centre National de la Propriété Forestière

GAM : Generalized Additive Model

GBM : Generalized Boosted Model

GIEC : Groupe d'experts Intergouvernemental sur l'Évolution du Climat

GLM : Generalized Linear Model

GRECO : Grande Région Écologique

IBBR : Institute for Bioscience and Biotechnology Research

IFN : Inventaire Forestier National

INRAE : Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

ONF : Office National des Forêts

RCP : Representative Concentration Pathway

RF : Random Forest

ROC : Receiver Operating Characteristic

SDM : Species Distribution Model

SER : Sylvoécocorégion

SIG : Système d'Information Géographique

SRC : Species Range Change

TSS : True Skill Statistic

WWF : World Wide Fund for Nature

Introduction

Les forêts sont directement concernées par le dérèglement climatique. En effet, avec l'augmentation globale des températures moyennes, viennent les effets secondaires d'un tel dérèglement : feux de forêts, vagues d'insectes, sécheresses, pathogènes... (Seidl et al., 2017), autant de perturbations qui remettent en cause la longévité et la pérennité des peuplements forestiers. Ces perturbations ont un impact sur les services écosystémiques prodigués à nos sociétés par les écosystèmes forestiers (Thom & Seidl, 2016), qui risquent d'être largement diminués en même temps que la forêt perd en diversité. Face à une augmentation du régime de perturbations, les populations d'arbres forestiers peuvent répondre de différentes manières : modifier leur aire de répartition en colonisant de nouvelles niches correspondant à leurs anciens habitats grâce à la migration, adapter leur phénotype aux nouvelles conditions environnementales, ou disparaître (Aitken et al., 2008). Ces changements apparaissant de manière rapide, la migration et l'adaptation semblent être un défi de taille pour des populations d'arbres forestiers, qui ont parfois des générations très longues (Alberto et al., 2013). Cependant, ces organismes sont caractérisés par des flux de gènes importants, des populations aux tailles conséquentes et souvent une fécondité élevée, ce qui leur procure un potentiel d'adaptation non négligeable face aux changements globaux rapides.

L'adaptation des forêts au changement climatique est au cœur de la recherche en sciences forestières. C'est le cas de l'unité « écologie des forêts méditerranéennes » (URFM) du centre de recherches INRAE PACA, à Avignon (84). L'URFM s'attache à comprendre, au travers d'angles de recherche pluridisciplinaires, la dynamique, le fonctionnement et l'évolution des forêts méditerranéennes. Cela passe par la compréhension de ces phénomènes au travers de plusieurs disciplines : écologie fonctionnelle, biologie évolutive, biogéographie, génétique, sciences forestières, mais aussi de divers outils : modélisation, expérimentation... Les enjeux soulevés sont autant fondamentaux qu'appliqués et concernent les milieux naturels comme les exploitations forestières. L'URFM intègre une approche prédictive au regard des changements globaux et vise notamment à prédire les risques liés au dérèglement climatique, dans une optique de stratégie de gestion de la ressource forestière.

Les problématiques auxquelles vont faire face les forêts françaises sont également le cœur du projet MedForFutur (CRPF, s.d.). Ce projet lancé par le CNPF, est financé par le ministère de l'Agriculture et de l'Alimentation dans le cadre de l'appel à projet « Innovation et investissements pour l'amont forestier ». Il regroupe l'ONF et INRAE, et a pour objectif l'adaptation de la gestion forestière méditerranéenne au contexte climatique. Cette adaptation passe notamment par la diversification génétique des forêts. En effet, il a été montré que la diversité génétique au sein d'un massif est à l'origine d'une plus grande résilience face à des menaces liées au réchauffement global (par exemple (Vilà-Cabrera et al., 2018)). Le projet vise donc à adapter les pratiques forestières face à l'appauvrissement constaté de la forêt méditerranéenne, les pins sylvestres (*Pinus sylvestris*) et les chênes pubescent (*Quercus pubescens*) étant les plus fortement touchés actuellement (Bouillie, ONF, 2021). Par ailleurs, les essences de la région méditerranéenne sont des ressources génétiques potentielles pour la diversification des forêts françaises. Face à la prochaine « méditerranéisation » (Barredo et al., 2018) du climat en France, les différentes espèces provenant de la région méditerranéenne représentent un levier intéressant pour endiguer un appauvrissement génétique des forêts plus continentales faisant face aux effets secondaires dévastateur du changement climatique. La question de l'adaptation des forêts est donc primordiale pour envisager ce que sera la forêt

française de demain, et l'étude de cette adaptation, qu'elle soit naturelle ou induite par nos sociétés est plus que jamais un point-clé dans la gestion de la ressource forestière.

Le projet MedForFutur a notamment pour vocation de soutenir les travaux de recherche allant dans cette voie, et c'est dans ce contexte institutionnel que ma mission de stage s'inscrit. En effet, ce travail de stage repose sur l'étude du réseau de dispositifs expérimentaux en jardins communs géré par INRAE au niveau national : le réseau GEN4X (Fady, 2018). L'objectif du stage était d'évaluer la représentativité et les potentialités de ce réseau dans les climats français actuel et futurs. Comprendre la représentativité climatique d'un tel réseau est un enjeu d'importance pour la recherche et la filière forestières, par exemple pour cibler les dispositifs expérimentaux qui sont actuellement présents dans les climats futurs ou ceux qui sont dans des habitats climatiques qui n'existeront plus. Dans la mesure où ce genre de dispositif expérimental est amené à donner des résultats sur une longue durée, et qu'ils ne sont pas délocalisables (plantations d'arbres), la prédiction des zones à enjeux sur un tel réseau est intéressante pour l'étude de l'adaptation des essences se trouvant dans les jardins communs. Ainsi, dans un premier temps, je me suis attaché à décrire en quoi les dispositifs en jardins communs permettent de comprendre l'adaptation des espèces forestières en testant leur diversité génotypique et phénotypique, afin de cibler les enjeux auxquels ils font face dans un contexte de changement climatique. Après cet état de l'art, il a été nécessaire de corriger et compléter la base de données du réseau GEN4X pour avoir une bonne prise en main des données. Enfin, grâce à des outils de modélisation de niche écologique, j'ai calculé l'habitat climatique du réseau des dispositifs en jardins communs sous différents scénarios, et mis en perspective ces résultats avec la gestion du réseau.

Etat de l'art

1. Contexte scientifique

1.1. Quelques éléments d'analyse bibliométrique

L'objectif de cette partie est de caractériser l'intérêt de la communauté scientifique pour les expérimentations en jardins communs, sous le prisme de l'adaptation locale, la plasticité phénotypique mais aussi la sélection variétale.

1.1.1. Méthode

Les équations de recherche par sous-thématique utilisées dans le moteur de recherche avancée Web of Science ont servi de bases pour cadrer les sujets en termes de publications. Ce moteur de recherche a été privilégié pour les possibilités de classement des résultats selon divers filtres et facteurs et les statistiques permises par le moteur. Des termes ont été ajoutés à ces équations afin de faire ressortir de potentielles tendances dans l'évolution du nombre de publications et l'implication des pays dans ces domaines de recherche. Ainsi, les 3 sous-thématiques ont été recherchées avec ou sans le facteur « forêt ». Les sujets de l'adaptation locale et de la plasticité phénotypique ont également été testés séparément avec le sujet des jardins communs. Une autre sous-thématique a été explorée : l'amélioration et la sélection des arbres forestiers en prenant en compte ou non les dispositifs en jardins communs. Enfin, la requête sur les jardins communs a été affinée selon des catégories se référant aux biomes (WWF, s.d.) : « boreal/taiga », « temperate », « subtropical », « mediterranean » et « tropical ». Les requêtes ont ensuite été analysées grâce à l'outil « analyze results », en termes d'évolution temporelles des publications, de pays les plus producteurs d'articles, et d'écosystèmes concernés par les publications. Les données concernant l'année 2021 sont exclues.

1.1.2. Tendances temporelles

En ce qui concerne les expérimentations en jardins communs, on en trouve quelques occurrences dès 1966 sur Web of Science (dont le contenu date au plus tôt de 1955). La requête affinée avec le domaine des sciences forestières montre que les premières recherches concernant ces dispositifs datent de 1991. Ce résultat est inattendu compte tenu de Langlet, 1971. En effet, comme il a été vu dans l'introduction, les expérimentations impliquant des plantations comparatives avec des arbres forestiers notamment existent depuis le XIX^{ème} siècle. Il est possible que les premières références de ces dispositifs existent surtout dans des revues spécialisées telles que *Silvae Genetica*, ou même dans la littérature grise. Une exploration plus fine montre que les premières occurrences du terme en écologie végétale datent de 1981, mais ne concernent pas encore les arbres forestiers. De manière générale, une augmentation de l'intérêt pour les dispositifs en jardins communs est notable, puisque le nombre de publications sur le sujet a été multiplié par 12 de 2000 à 2020. L'intérêt pour ces dispositifs en sciences forestières se manifeste d'autant plus avec une multiplication du nombre de publications par 17 sur la même période.

Concernant l'adaptation locale, une augmentation similaire est observée. C'est en 1977 qu'un premier article emploie ce terme à des fins naturalistes, et on remarque que son emploi est multiplié par 12 depuis 2000. En précisant la requête à la diversité génétique, on remarque également une augmentation du nombre de publications, qui est multiplié par 16 en 20 ans. Les

premières études trouvées avec l'équation de recherche sont cependant plus récentes (1990). Ces sujets appliqués aux sciences forestières font également l'objet d'un intérêt croissant, avec une multiplication par 38 du nombre de publications depuis 2000, en notant une forte croissance de ce nombre depuis 2010.

Le sujet de la plasticité phénotypique connaît une augmentation de l'intérêt plus douce, le nombre de publications depuis 2000 étant multiplié par 5. Cela se ressent également si l'on considère le sujet appliqué aux sciences forestières, avec une multiplication du nombre de publication par environ 4 ces 20 dernières années. L'intérêt pour ces thématiques semble avoir débuté en même temps que pour l'adaptation locale et la diversité génétique.

Enfin, sous un autre angle d'analyse, le sujet de l'amélioration génétique forestière est considéré. On constate un intérêt faible mais présent dès les années 70, qui ne connaît une réelle augmentation que dans les années 90. Le nombre de publications a été multiplié par 6 en 20 ans. Si l'on ajoute à cette thématique celle du jardin commun, on constate un intérêt plus récent (1992), avec une croissance globale du nombre de publications mais une croissance variable d'années en années.

Deux autres croisements ont été effectués. Les sujets de l'adaptation locale et de la plasticité phénotypique ont été recherchés avec la thématique des jardins communs, afin de constater l'évolution de l'importance de ces dispositifs dans la littérature. En ce qui concerne l'adaptation locale et la diversité génétique, le nombre de publications qui évoquent les dispositifs en jardins communs a été multiplié par 35 en 20 ans, avec une première publication en 1994. Concernant la plasticité phénotypique, le nombre de publications évoquant les jardins communs a été multiplié par 8 depuis 2000. On remarque donc une tendance à évoquer et utiliser ces dispositifs expérimentaux davantage pour parler d'adaptation locale et diversité génétique que de plasticité phénotypique (quoique la différence du nombre de publications totales soit relativement faible, 375 contre 201).

On note donc dans tous les cas un intérêt croissant pour ces thématiques dans le monde de la recherche, avec souvent une augmentation plus nette des publications ces vingt dernières années. Cette augmentation peut en partie s'expliquer par un intérêt croissant de la communauté scientifique pour les sujets concernant les réponses environnementales aux changements globaux actuels, sujets bien représentés dans les thématiques d'intérêt de cette revue. A titre de comparaison, la requête simple « forestry » (filtrée ensuite selon la catégorie Web of Science « forestry ») montre une augmentation des publications depuis les premières occurrences du terme dans des articles scientifiques et techniques (1965), mais cette augmentation semble plus modérée, puisque le nombre de publications a seulement été multiplié par 2 depuis 2000. De la même manière, la requête « ecology » (filtrée ensuite selon la catégorie Web of Science « ecology ») montre sans grande surprise un intérêt antérieur pour le sujet (1955 voire avant), et un nombre de publications multiplié par plus de 3 de 2000 à 2020. Les thématiques abordées dans ce mémoire semblent donc faire l'objet d'une forte augmentation dans la publication d'articles ces dernières années.

1.1.3. Implication des pays

Pour les requêtes « jardins communs », « adaptation locale », « plasticité phénotypique », avec ou sans le mot-clé « forêt » et « amélioration génétique » avec ou sans le mot-clé « jardin commun », une liste des publications par pays a été réalisée. Le pays le plus impliqué dans les publications toutes requêtes confondues (et donc toutes thématiques confondues) sont les Etats-Unis, qui concentre systématiquement entre 23 et 42% des publications totales par requête. On trouve ensuite 12 autres pays qui sont systématiquement présents dans la liste des 20 premiers pays qui publient par requête (non classés) : le Canada, l'Allemagne, la France, l'Espagne, la Chine, l'Australie, la Suisse, la Suède, la Finlande, l'Italie, le Danemark et le Royaume-Uni. Ce résultat est peu surprenant, au regard de leur

implication dans la recherche scientifique mondiale (OST, 2021). En effet, en 2015 on retrouve parmi les 20 premiers pays en termes de nombre de publications : les Etats-Unis, la Chine, le Royaume-Uni, l'Allemagne, la France, l'Italie, le Canada, l'Espagne, l'Australie et la Suisse. On note que les pays les plus impliqués dans ces thématiques sont des pays du Nord, possédant également une importante ressource forestière (les pays européens cités sont parmi ceux possédant la plus grande couverture forestière du continent), mais qui possèdent des forêts méditerranéennes, tempérées et boréales. D'autres pays moins présents dans les listes des 20 premiers pays productifs par requête sont toutefois impliqués dans la recherche : la Pologne, la Belgique, les Pays-Bas, le Mexique, l'Argentine et le Brésil. On voit apparaître des pays au développement plus récent, qui couvrent par ailleurs des zones climatiques plus méridionales, avec des types de forêt subtropicaux à tropicaux. Ainsi, ces thématiques de recherche concernent surtout des pays développés de l'hémisphère nord et laissent penser que les expérimentations en jardins communs concernent majoritairement des forêts boréales à méditerranéennes.

1.1.4. Ecosystèmes concernés

La requête sur les jardins communs et les forêts (tout type) donne 1055 résultats. D'après les catégories Web of Science, ces résultats semblent plutôt concerner le sujet d'intérêt (ecology, forestry, plant science, evolutionary biology...). Ainsi, en ajoutant aux requêtes les mots-clés qui réfèrent aux biomes, on obtient les résultats suivants. La requête « boreal/taiga » donne 49 résultats. La requête « temperate » donne 154 résultats. La requête « mediterranean » donne 56 résultats. La requête « subtropical » donne 13 résultats. Enfin, la requête « tropical » donne 84 résultats. Les requêtes par grands écosystèmes sont loin de couvrir l'ensemble des résultats trouvés au départ : 699 références ne seraient pas répertoriées selon ces mots-clés « biomes » dans l'hypothèse où il n'existerait pas d'intersection entre les requêtes formulées. La requête « jardins communs et forêts » a donc été saisie en excluant les mots clés propres aux écosystèmes, 831 résultats sont trouvés, ce qui semble cohérent. Etant dans l'incapacité de consulter l'ensemble de ces références, on émet l'hypothèse que ces 831 références n'évoquent pas explicitement les biomes formulés précédemment, ou même abordent des concepts plus théoriques. Bien que cette méthode d'analyse par requête semble avoir d'importantes limites en termes d'exhaustivité, on note tout de même que dans les résultats trouvés par biomes, ce sont les forêts tempérées qui sont le plus représentées dans les références à propos des jardins communs. Ce résultat concorde avec l'idée formulée dans la partie précédente selon laquelle les forêts les plus étudiées seraient celles présentes dans les pays les plus productifs en termes de littérature scientifique sur ces thématiques, i.e. les pays développés du Nord. Le second grand écosystème représenté serait la forêt tropicale, qui gagnerait à être étudié davantage (Savolainen et al., 2007).

1.2. Les dispositifs en jardins communs

Les peuplements forestiers doivent faire face à de nouvelles problématiques, conséquences des changements climatiques. L'adaptation des peuplements forestiers est plus que jamais un enjeu pour leur résilience et pour une gestion durable de la ressource. Un bon moyen d'étudier les phénomènes d'adaptation est un dispositif expérimental appelé « jardins communs ». Cet outil sert à mettre en évidence une certaine diversité génétique entre individus mais aussi à prouver l'existence de plasticité phénotypique. On en trouve les premières traces dès le XVIIIème siècle, à l'époque où l'inspecteur-général de la Marine française et botaniste, Duhamel du Monceau a imaginé les premiers tests de provenance sur des cultures de *Pinus sylvestris*, dans l'optique d'augmenter les stocks de bois pour la construction de navires

(Langlet, 1971). C'est dans les années 1820 que les tests de provenance sont réellement étudiés grâce à De Vilmorin, qui reprend les questionnements de Duhamel, mais cette fois-ci publie les résultats en 1862. S'ensuit un intérêt croissant pour les tests de provenance, si bien qu'en 1892 l'IUFRO (Union of Forest Research Institution) est créée (Matyas, 1996). Parmi ses objectifs premiers, on retrouve l'initiation et la coordination de tests de provenance internationaux. C'est le début de ce que l'on appelle la « genecology », et qui donne ensuite lieu à de plus en plus d'études sur les liens entre provenances, généalogie, climats, sols, phénotypes, dans une optique d'amélioration et d'augmentation de la ressource forestière, mais aussi de connaissance fondamentale, avec l'étude des phénomènes d'adaptation. Les jardins communs sont donc des dispositifs anciennement utilisés, mais qui semblent aujourd'hui toujours aussi utiles et efficaces.

On appelle « jardin commun » (« common garden experiment »), un dispositif expérimental dans lequel on introduit des individus ou génotypes dans un environnement commun plus ou moins contrôlé afin d'observer comparativement leur croissance (Kawecki & Ebert, 2004; Lefèvre et al., 2015; De Villemereuil, 2016). Un tel dispositif permet de tester l'origine génétique des différences observées entre individus dans la plantation comparative. En effet, puisque tous les individus sont cultivés dans le même environnement, on considère que la variabilité environnementale qui entre dans la variation phénotypique observée est nulle. En fonction des tests, des caractères observés et des unités génétiques, on observe une plus ou moins forte différenciation entre individus, directement imputables à de la diversité génotypique et indice d'une potentielle adaptation locale. Les unités génétiques testées peuvent être des clones, des familles (on connaît leurs ascendants) ou encore des provenances (on connaît leur aire géographique d'origine) (Lefèvre et al., 2015).

Une expérimentation contenant plusieurs unités génétiques dans un environnement commun permet de comprendre les variations phénotypiques imputables à une diversité génotypique, et ce, sans passer par des techniques de génomique. Il est également possible de détecter la variabilité phénotypique due à l'effet de l'environnement, autrement dit à une certaine plasticité, en répliquant des expérimentations en jardins communs dans des environnements différents (Alberto et al., 2013). En effet, si on observe des différences phénotypiques (adaptatives ou non) entre différents environnements au sein d'une même unité génétique (des clones dans l'idéal), on peut conclure à une plasticité phénotypique due à une influence systématique de l'environnement. Au contraire, une absence de différence phénotypique entre les différents environnements suggère une absence de plasticité (Lefèvre et al., 2015), au moins dans les gammes d'environnement testées.

Les dispositifs en jardins communs restent donc un outil puissant pour mesurer la diversité génétique et apprécier la variabilité phénotypique de populations forestières. Si ces dispositifs permettent parfois de suggérer un phénomène d'adaptation locale, ils ne sont pas encore suffisants pour émettre des conclusions à ce sujet. C'est grâce à des dispositifs particuliers appelés transplantation réciproque que l'adaptation locale peut-être prouvée.

1.3. Méthodologie du jardin commun forestier

Après avoir considéré les aspects théoriques des expérimentations en jardins communs, nous allons voir quelques indispensables considérations pratiques. En effet, ces dispositifs expérimentaux nécessitent l'élaboration de protocoles robustes qui permettent ensuite d'effectuer des mesures exploitables, notamment en sciences forestières. Cette partie est largement inspirée des travaux de Wright dans « Introduction to Forest Genetics » (1976), qui restent très actuels.

Tout d'abord, le principe de base d'un jardin commun est l'absence de variabilité environnementale. Tout doit être fait pour que le milieu soit uniforme en tout point. De cette manière, la composante environnementale de la variabilité phénotypique peut être écartée. Il est bien sûr impossible d'obtenir un milieu parfaitement uniforme, c'est pourquoi il est possible de considérer les variations micro-locales dans un modèle statistique (Lefèvre et al., 2015). Plusieurs designs peuvent être adoptés pour une telle expérimentation mais le plus courant reste peut-être le « Randomized complete block design » qui consiste en une plantation divisée en un nombre de blocks de taille équivalente, chacun contenant un plot de chaque seedlot. On applique une « randomization » qui consiste à distribuer les seedlots de manière aléatoire dans un block, afin d'éviter les biais statistiques liés à un milieu potentiellement non uniforme en ressources au sein des blocks.

Une fois le design de l'expérimentation effectué, plusieurs aspects pratiques sont à prendre en considération car ils risquent d'influer sur les résultats finaux : la qualité des semences, la sélection du site de plantation, l'espacement, l'alignement, l'étiquetage et la cartographie des sites de plantation, le soin porté à la protection de la plantation (adventices, ravageurs...). Des considérations statistiques sont également à prendre en compte : la taille des seedlots, des plots, sont à étudier pour que les analyses statistiques soient significatives. Il est dès lors nécessaire de discuter de valeurs seuils minimales d'échantillons afin d'obtenir des résultats précis.

Les dispositifs en jardins communs permettent de nombreux tests, impliquant des relations entre seedlots différentes : les géotypes testés peuvent être des clones, auquel cas on s'attend à très peu de variabilité génétique entre individus, ils peuvent être également reliés par des liens de parenté (half-sibs/full-sibs) ou encore être des descendants d'individus issus d'une même aire géographique (Lefèvre et al., 2015). Dans ce dernier cas, on appelle les plantations comparatives des tests de provenance. Ces tests peuvent avoir des objectifs différents. En effet, on retrouve des expérimentations à portée théorique, destinées à produire de la connaissance fondamentale, notamment sur des phénomènes évolutifs. D'autres expérimentations ont une portée plus appliquée, comme l'étaient d'ailleurs les premiers tests de provenance (Langlet, 1971), dans une optique de rendement forestier, de sélection variétale ou d'amélioration génétique.

Les étapes d'une expérimentation en jardin commun peuvent varier en fonction des protocoles mais possèdent des similitudes. La collecte des graines nécessaires à l'expérimentation s'effectue en fonction des objectifs de l'étude. Si on s'intéresse à des variations de traits dues à des différences climatiques, on peut collecter des provenances ou lots de graines selon des gradients latitudinaux (Matyas, 1996), longitudinaux (Martínez-Sancho et al., 2021), ou altitudinaux (Oleksyn et al., 1998; Vitasse et al., 2009; Castellanos-Acuña et al., 2015). Il arrive également que des graines soient collectées en fonction du substrat d'origine des provenances (Cordell et al., 1998). Il est ensuite courant de faire pousser les graines dans des pépinières, en laboratoire ou sous serres, en conditions très contrôlées. Les étapes de germination et de début de croissance s'y déroulent. Les seedlots sont ensuite généralement

Glossaire « Méthodologie du jardin commun forestier » (d'après Wright, 1976)

Seedlot : groupe d'individus partageant un lien (apparentés, clones...) identifiés comme unité de base du jardin commun

Plot : groupe d'individus appartenant à un unique seedlot plantés de manière adjacente

Block : portion du jardin commun contenant 1 plot de chaque seedlot considéré

Nursery : plantation particulière dans laquelle les seedlot passent le début de leur vie (germination, croissance) dans l'attente d'être plantés dans un jardin commun définitif

Replication : plantation de différents plots d'un même seedlot à plusieurs endroits au sein d'une plantation, dans différentes plantations ou à des périodes différentes

Half-sibs/Full-sibs : individus avec un parent en commun (half) ou deux parents en commun (full)

Provenance : origine géographique d'un arbre

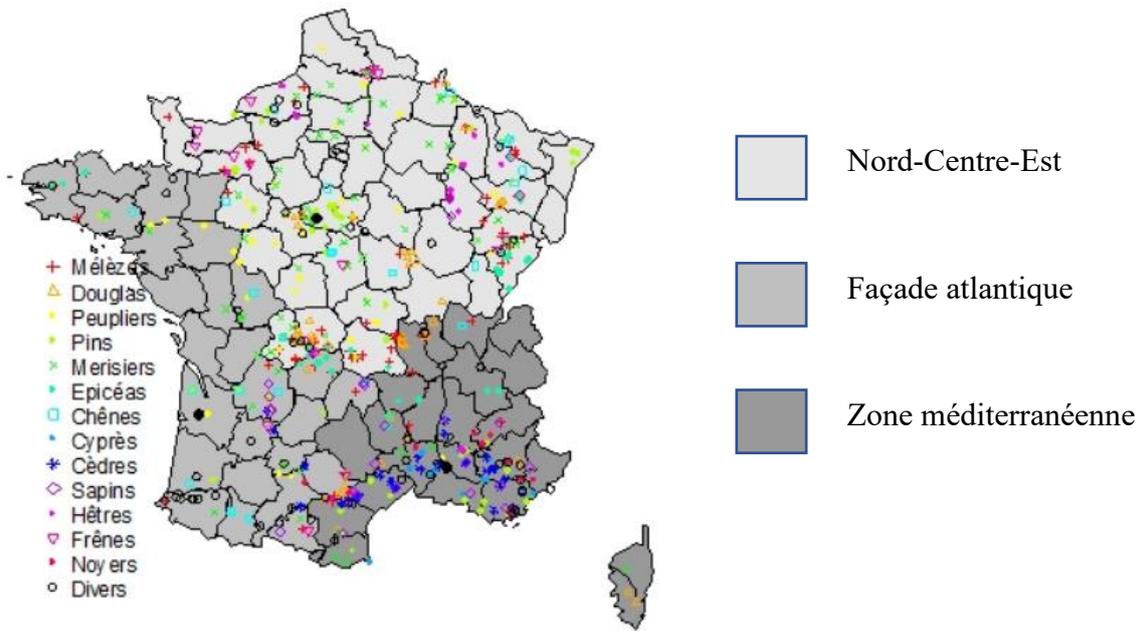


Figure 1 : Localisation des différents dispositifs GEN4X et espèces, en fonction des 3 zones de gestion du réseau (source : site web de IN-SYLVA)



Figure 2 : Photographies d'un dispositif en jardin commun du réseau GEN4X : Test de provenances de sapin de Céphalonie (*Abies cephalonica*) situé à Saint-Lambert (84)

transférés définitivement dans des dispositifs en jardins communs dans des « field sites », dont les conditions environnementales sont plus ou moins contrôlées et selon les dispositions nécessaires à l'expérimentation (taille des échantillons, disposition des blocks, localisation du site et conditions pédoclimatiques...). Il arrive que les expérimentations soient répétées à différentes localisations, à des altitudes différentes par exemple (Castellanos-Acuña et al., 2015). Des mesures de morphologie, de physiologie, de phénologie et de variables environnementales peuvent ensuite être effectuées avant d'analyser statistiquement ces résultats.

1.4. Les Species Distribution Models, un outil puissant

La modélisation de niches écologiques fait l'objet de nombreuses publications. L'efficacité de ces modèles pour connaître et prédire l'aire de répartition d'une ou de plusieurs espèces représente un des grands enjeux de la recherche en écologie, dans un contexte de changements globaux rapides. Les SDMs (Species Distribution Models) sont aujourd'hui largement répandus au sein de la littérature scientifique dans les domaines de l'écologie et de la biologie évolutive (Araújo & Guisan, 2006). On trouve deux types de modèles. Les modèles mécanistes se basent sur la physiologie de l'espèce pour établir un lien entre les traits fonctionnels et leur environnement (Kearney & Porter, 2009). Certains considèrent néanmoins que ces modèles ne rentrent pas dans la catégorie des SDMs (Jane Elith & Leathwick, 2009). Les modèles corrélatifs quant à eux permettent à partir de données de présence avérée (distribution de l'espèce), d'extrapoler spatialement cette distribution en utilisant des outils de régression statistique, et ce, afin de modéliser une aire de répartition potentielle (Jane Elith & Leathwick, 2009). Cette extrapolation, en plus d'être spatiale, peut également être temporelle en s'appuyant sur des scénarios environnementaux futurs. Ce type de modèle représente environ 90% des publications sur les modèles de niches (Pecchi et al., 2019). Les points clés de la modélisation de niches sont les suivants : la collecte de données pertinentes et précises, la sélection d'algorithmes, le calibrage du modèle avec des données d'entraînement, l'évaluation des modèles au moyen d'indicateurs et l'amélioration des modèles en répétant le processus. Bien que très efficaces aujourd'hui, l'amélioration des SDMs passe par une meilleure réalisation de ces points-clés (Araújo & Guisan, 2006). Les SDMs sont de plus en plus mobilisés en sciences forestières, grâce à l'augmentation des jeux de données disponibles (Benito-Garzón & Fernández-Manjarrés, 2015; Pecchi et al., 2019; Ray et al., 2021). Ils sont aujourd'hui utilisés comme un outil fiable pour prédire les aires de répartition future des espèces forestières qui font face au changement climatique, et ainsi organiser une gestion durable de cette ressource vulnérable.

2. Le réseau GEN4X

Le réseau GEN4X (réseau de GENétique FOrestière pour la Recherche et l'eXpérimentation) regroupe l'ensemble des dispositifs expérimentaux en jardins communs, géré par le département ECODIV (écologie et biodiversité) d'INRAE. Le réseau regroupe donc l'ensemble des tests effectués sur des unités génétiques (espèce, provenance, descendance, clones...) d'essences forestières (figure 1 et 2). La volonté de regrouper l'ensemble de ces dispositifs expérimentaux a vu le jour en 2009, avec l'émergence d'un premier réseau nommé « PlantaComp » qui regroupait déjà une partie des plantations comparatives. C'est en 2018 que le réseau GEN4X est officiellement créé et regroupe l'ensemble de ces dispositifs expérimentaux (plus d'un millier), dans un triple objectif scientifique, d'animation d'une communauté technique et de centralisation des données de mesures de terrain. GEN4X est un des réseaux de l'infrastructure de recherches nationale IN-Sylva France. Le réseau comprend plus de 2000 ha de tests pour un total d'environ 2,5 millions d'arbres plantés. Il sert aujourd'hui

à l'expérimentation, à l'observation mais aussi à la conservation de ressources génétiques forestières. De nombreuses mesures sont effectuées dans les dispositifs : phénologie, croissance (mesures de hauteur, croissance radiale), survie (résistance aux parasites, maladies), fructification, mais aussi des récoltes d'échantillons biologiques indispensables aux analyses de biologie moléculaire (feuilles, graines). Le réseau produit un grand nombre de données et métadonnées qui sont gérés par les unités expérimentales INRAE associées. Le réseau est en effet partitionné en 3 aires géographiques : la zone méditerranéenne gérée par l'UEFM à Avignon, la façade atlantique gérée par l'UEFP à Bordeaux, et le Nord-Centre-Est géré par GBFOR à Orléans. Les unités expérimentales travaillent en lien avec les chercheurs du département ECODIV, qui ont la plupart du temps plusieurs dispositifs à leur charge où se réalisent des expérimentations. Les axes de recherches qui caractérisent le réseau sont aujourd'hui tournés vers l'adaptation des peuplements forestiers au dérèglement climatique et toutes les conséquences qui en découlent : dépérissement, appauvrissement de la diversité génétique, risques sanitaires... C'est pourquoi cet outil s'avère très important dans l'étude de tels changements. Un réseau de plus d'un millier de jardins communs documentés est une importante source d'information pour la prédiction des risques auxquels vont faire face les forêts françaises, et représente un outil de taille pour comprendre l'adaptation de différentes espèces et unités génétiques à de nouveaux climats et permettre une création variétale résiliente. Afin d'orienter de nouveaux projets de recherche concernant la résilience et la diversification des forêts françaises au moyen de l'étude de l'adaptation grâce aux jardins communs, il est d'abord nécessaire de connaître la représentativité climatique actuelle du réseau et sa couverture climatique future.

C'est dans un tel contexte que ma mission de stage s'inscrit. Connaître les climats auxquels vont faire face les dispositifs, et savoir à quel point le réseau est susceptible de gagner ou de perdre en représentativité et diversité climatiques représente un enjeu de gestion et de prospective scientifique. Dans ce cadre de changements manifestes, il est intéressant de s'intéresser à une partie du réseau dont le climat est susceptible de couvrir une plus grande partie du territoire français à l'horizon 2050 : la région méditerranéenne.

Concrètement, mon stage répond à la problématique suivante :

Dans quelle mesure est-il possible de déterminer l'aire de compatibilité climatique actuelle et future du réseau GEN4X ? Quelle est la représentativité du réseau GEN4X en termes de climats actuels et futurs ? Dans un contexte d'adaptation des forêts françaises, quelle est la place des dispositifs expérimentaux méditerranéens dans cette représentativité climatique ?

On suppose en premier lieu qu'en utilisant des outils de modélisation type « Species Distribution Models », il sera possible d'obtenir une prédiction robuste de la niche climatique du réseau GEN4X. Ensuite, on suppose que le réseau, de par son ampleur, représente une grande partie du climat français, mais que cette représentativité va être largement modifiée à l'avenir. Enfin, il est possible que les dispositifs méditerranéens prennent une place de plus en plus importante dans l'étude de l'adaptation des peuplements, dans une optique de diversification des forêts françaises.

Matériel et méthodes

1. Aire d'étude

Le réseau GEN4X est réparti sur l'ensemble du territoire français métropolitain (dont la Corse). L'étude suivante a donc été menée sur ce territoire, exclusivement. Une attention particulière a été portée à la région climatique méditerranéenne.

2. Partenaires et acteurs

Ce travail de stage a été réalisé sous l'impulsion et la direction de Bruno Fady, directeur de recherche à l'URFM (écologie des forêts méditerranéennes) à INRAE Provence Alpes Côte d'Azur. Il a été mené avec l'appui de Denis Vauthier (Unité expérimentale Entomologie et Forêts Méditerranéenne, INRAE PACA) mais aussi Marta Benito Garzón (UMR Biogeco, INRAE Bordeaux), qui a apporté son expertise. Des données nécessaires pour compléter la base de données ont été transmises par Frédéric Bernier (Unité Expérimentale Forêt Pierroton, INRAE Bordeaux) ainsi que Dominique Veisse (Unité Expérimentale GBFOR, INRAE Orléans).

Les données environnementales ont été transmises par Maurizio Marchi, chercheur au Institute of Biosciences and BioResources à Florence. Il a également apporté son appui en termes de conseil sur l'analyse des données.

Enfin, Wilfried Thuiller, directeur de recherche CNRS au Laboratoire d'écologie alpine (Grenoble) et notamment co-développeur du package *biomod2*, a également été sollicité pour des conseils d'analyse et d'utilisation du modèle.

3. Données de présence des dispositifs

Ce travail repose sur la base de données du réseau GEN4X. Cette base se compose de plusieurs entrées qui correspondent aux caractéristiques des dispositifs expérimentaux du réseau. On y trouve, pour chaque dispositif (c'est-à-dire jardin commun), de nombreuses métadonnées : des informations sur le site expérimental (codification, espèce, type de test), des informations de gestion internes au laboratoire, des données de localisation (coordonnées, géographie du site), des relevés écologiques ainsi que des données techniques de composition des sites. La base telle qu'elle était en début de stage comportait 1213 dispositifs (dont 70 dispositifs avec des coordonnées manquantes), répartis sur l'ensemble du territoire français. Le stage a par ailleurs été l'occasion de compléter et de corriger la base. Pour l'analyse, les informations utiles à l'analyse sont essentiellement l'index (code du dispositif dans la base), les espèces testées dans les dispositifs, la localisation en coordonnées décimales.

3.1. Préparation des données

Un travail préliminaire de complémentation de la base a été nécessaire afin d'obtenir les données utiles à l'analyse. En effet, certains dispositifs gérés par UEFP et GBFOR ne se trouvaient pas dans la base d'origine. C'est pour cela que Dominique Veisse et Frédéric Bernier

ont été sollicités afin de remédier à ces manques. Un travail de complémentarité a également été effectué à partir de la base gérée par l'UEFP. Dans un second temps, un contrôle des dispositifs présents en plusieurs exemplaires a été effectué. Dans la perspective d'une analyse spatiale et lors de la création des jeux de données, les dispositifs aux localisations aberrantes ont été écartés et les coordonnées décimales manquantes ont été soit complétées, soit écartées en cas de manque d'information. Enfin, afin de faciliter la manipulation des données entre les différents supports utilisés (Microsoft Excel, QGis, RStudio), une codification arbitraire basée en partie sur les règles de codification des dispositifs a été établie pour les dispositifs aux codes manquants ou erronés.

3.2. Format des données

Les données exploitables utilisées dans ce travail comportent donc pour chaque dispositif : un code dispositif ainsi que des coordonnées décimales. Ces données ont permis la constitution de 3 jeux de données directement issus de la base de données GEN4X.

- Un premier jeu de données « FRANCE » représente l'ensemble du réseau GEN4X réparti sur le territoire métropolitain. Ce jeu de données compte 1208 jardins communs (les dispositifs aux localisations aberrantes ont été écartés) (figure 3).

Carte des dispositifs GEN4X répartis sur le territoire français (jeu de données FRANCE)

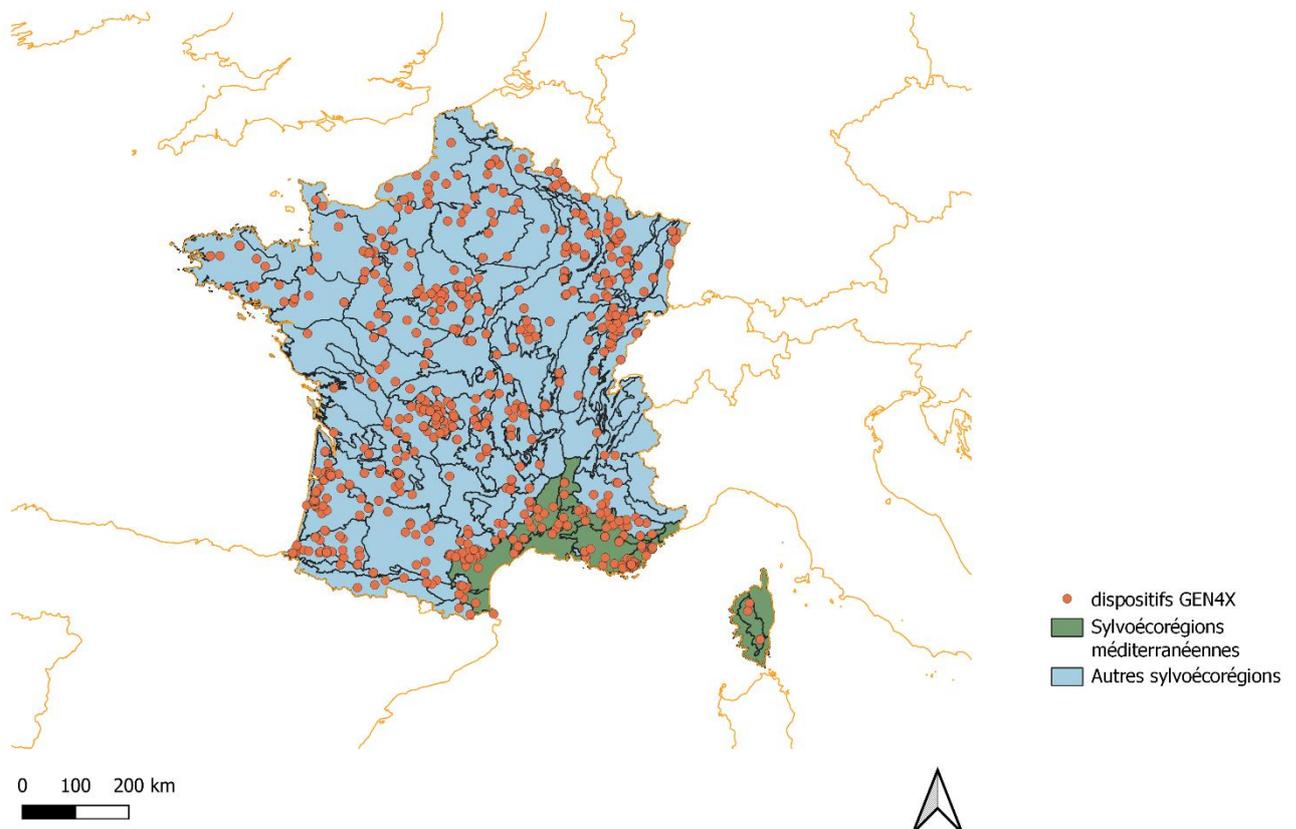


Figure 3 : Carte de répartition des 1208 jardins communs du réseau GEN4X en France métropolitaine (jeu de donnée FRANCE)

- Un second jeu de données « GRECO » représente les dispositifs dits « strictement méditerranéens ». Ce jeu de données a été constitué en ne retenant que les dispositifs GEN4X présents dans les Grandes Régions Ecologiques J et K (IGN, 2019), qui correspondent aux GRECO « Méditerranée » et « Corse ». Ces GRECO sont considérées ici comme abritant un climat méditerranéen. En effet, la correspondance au niveau géographique entre ces deux régions écologiques et l'écorégion « Mediterranean Forests, woodlands and shrubs » (WWF, s.d.) a permis de considérer les GRECO comme références en termes de délimitation du climat méditerranéen sur le territoire français. Ce jeu de données compte 181 individus (figure 4).

Carte des dispositifs GEN4X présents dans la zone climatique méditerranéenne (jeu de données GRECO)

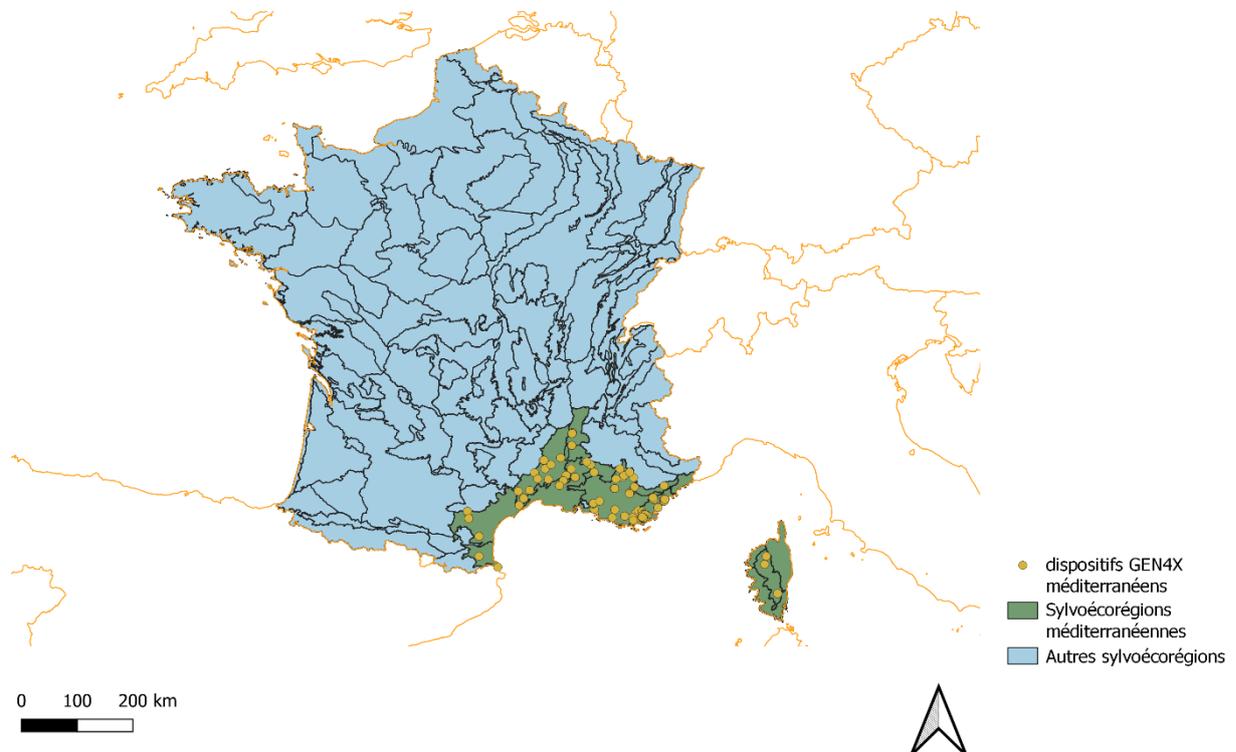


Figure 4 : Visualisation spatiale du jeu de données GRECO

- Un troisième jeu « ESP » de données représente les dispositifs qui testent au moins une espèce classée comme méditerranéenne. On considère ici qu'une espèce est méditerranéenne si elle possède la majorité de son aire de répartition en région méditerranéenne. Ainsi, ce jeu de données compte 185 jardins communs (figure 5).

Carte des dispositifs GEN4X contenant au moins 1 espèce méditerranéenne (jeu de données ESP)

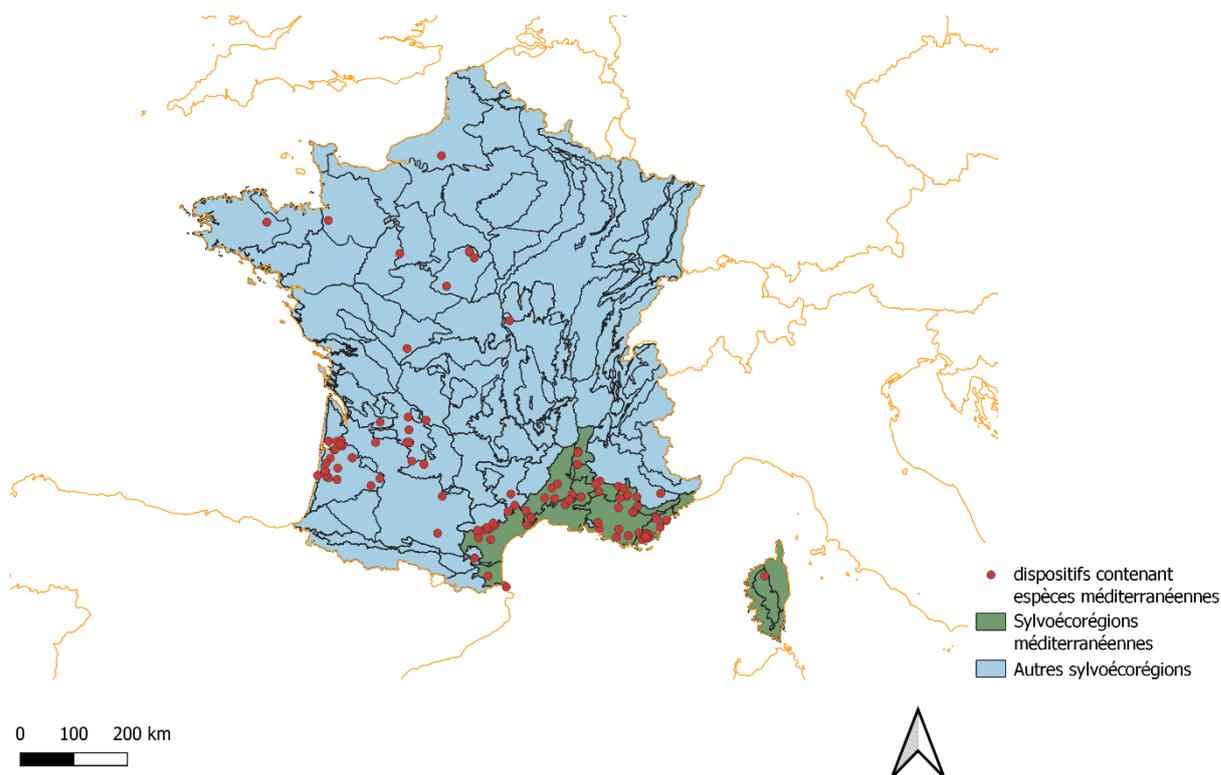


Figure 5 : Visualisation spatiale du jeu de données ESP

4. Données environnementales

Les données environnementales sont essentiellement climatiques. Elles sont issues de ClimateDT (ClimateDT, 2021), portail web qui fournit des données climatiques au niveau global en utilisant des données CRU-TS (Harris et al., 2020), pour la période historique (1901-actuel) et des données UKCP18 (Lowe et al., 2019) pour les projections futures (jusqu'à 2098). Ce système est une version en ligne des logiciels ClimateEU (Marchi et al., 2020), et ClimateNA, (Wang et al., 2016) et implémente une approche dynamique de downscaling « scale-free » pour fournir plus d'une centaine de variables et indices climatiques dédiés aux études écologiques, prédictions spatiales et pour tout travaux de recherches en général.

Les données issues de ClimateDT sont de deux types : des données spatiales sous format CSV utilisées pour l'analyse préliminaire à la modélisation et des données spatiales sous format Raster dédiées au travail de modélisation. Dans les deux cas, ces données ont été fournies par Maurizio Marchi (IBBR), développeur de l'outil, car les formats requis étaient trop importants pour passer par l'application web. Les données climatiques sous forme de rasters ont une précision d'1 km x 1 km et sont réparties sur l'emprise du territoire français (dont la Corse). Elles concernent 19 variables climatiques dérivées de valeurs de température et de précipitation. (Voir Annexes). Ces données ont été extraites pour des périodes de 31 ans : 1990-2020, 2050-2080 avec une projection climatique sous RCP 2.6 et 2050-2080 avec une projection climatique sous RCP 8.5. Les RCP (« Representative Concentration Pathways ») représentent des trajectoires de concentration de gaz à effet de serres établies par le GIEC dans son 5^{ème} rapport

d'évaluation (IPCC, 2014). Les RCP 2.6 et 8.5 représentent respectivement les scénarios les plus optimistes et pessimistes quant à la quantité future de gaz à effet de serre présente dans l'atmosphère, et permettent ainsi de modéliser le climat terrestre en conséquence. Ainsi, pour chaque période considérée, une moyenne sur 31 ans des valeurs annuelles des variables climatiques de *bio1* à *bio19* a été effectuée. On a donc 3 x 19 rasters issus de ClimateDT correspondant aux données climatiques utiles à cette étude.

5. Analyse descriptive et exploratoire

5.1. Description de la base

Une première analyse descriptive de la base de données GEN4X (après rajout de dispositifs) a été réalisée dans un objectif de caractérisation du réseau en termes d'espèces et de types de tests notamment. Cette analyse a été réalisée sur le logiciel de tableur Microsoft Excel. Une analyse spatiale préliminaire a également été effectuée au moyen du logiciel de SIG QGIS. Ce travail avait pour objectif de caractériser le réseau par rapport aux sylvoécotones françaises de l'IFN. La couche sous format shapefile des sylvoécotones a été extraite depuis le site *geo.data.gouv.fr* (Data.gouv.fr, 2010). L'objectif était de déterminer : le nombre de sylvoécotones représentées par des dispositifs expérimentaux, le nombre de dispositifs par sylvoécotone afin d'avoir un premier aperçu de la représentativité climatique du réseau GEN4X sur le territoire français.

5.2. Exploration des données climatiques

Une exploration du climat français a été réalisée dans un second temps. Pour cela, les données CSV issues de ClimateDT ont été utilisées. L'objectif était de comparer l'évolution dans le futur des valeurs climatiques du territoire français métropolitain aux valeurs climatiques actuelles du réseau GEN4X en mettant l'accent sur les dispositifs méditerranéens. 3 diagrammes ont été réalisés : un diagramme température/précipitation en utilisant les variables *bio1* et *bio12*, un diagramme mettant en avant les phénomènes de saisonnalité du climat en utilisant les variables *bio4* (saisonnalité de la température, en °C) et *bio15* (saisonnalité des précipitations, coefficient de variation), et un diagramme mettant en avant les phénomènes climatiques extrêmes en termes de température et de précipitation en utilisant les variables *bio11* (température moyenne du mois le plus froid, °C) et *bio14* (précipitations du mois le plus sec, mm). Les résultats sont visualisés sur RStudio grâce au package *ggplot*.

6. Modélisation des aires de compatibilité climatique du réseau GEN4X

6.1. Principe

Cette étude se base sur la modélisation de niches climatiques au moyen de Species Distribution Models » (SDMs). L'objectif de ce travail est de modéliser l'aire potentielle que pourraient occuper les dispositifs du réseau GEN4X en se basant sur des variables climatiques, au sein du territoire français. Ainsi, en s'appuyant sur les points de présence des dispositifs

expérimentaux, on simule la répartition spatiale d'une espèce. On extrapole ensuite la répartition potentielle du réseau en utilisant des SDMs. L'idée est donc d'utiliser les SDMs de manière « artificielle » (on ne travaille pas sur une espèce, on considère que le réseau GEN4X est une espèce), afin de modéliser la représentativité du réseau GEN4X en termes climatiques. C'est donc cela que l'on appelle ici la compatibilité climatique. Pour une espèce, cela représenterait l'habitat favorable potentiel où elle pourrait vivre. Ici, cela représente l'habitat représenté par les dispositifs GEN4X en termes climatiques.

6.2. Sélection des données environnementales

Tout d'abord, une sélection de variables climatiques parmi les 19 variables de ClimateDT a été effectuée, afin d'éviter tout phénomène de colinéarité entre les variables prédictives. Pour cela, une analyse en composante principale (ACP) a été réalisée sur les données climatiques *bio1* à *bio19* sur la période 1990-2020, qui sont les entrées du package *biomod2*. L'ACP a été réalisée moyen du package *Factominer*. La première étape a été de convertir les données climatiques en dataframe, ce qui correspond à une matrice de 933014 points, puisque chaque pixel des 19 couches rasters correspondant a été converti en données matricielles. Puis, l'analyse a pu être menée sur le jeu de données résultant. Les fonctionnalités de *Factominer* ont notamment permis d'afficher le graphique des valeurs propres, le cercle de corrélation, ainsi qu'un biplot faisant apparaître les graphes des variables (les données actives de l'ACP ainsi que les points des jeux de données concernés) et des individus. Cette analyse a ensuite permis de sélectionner un jeu de données plus petit mais qui représente au mieux les variations climatiques rencontrées sur le territoire français. Ainsi, on sélectionne des variables qui sont le plus indépendantes possibles, c'est-à-dire non colinéaires, car des phénomènes de colinéarité entre les variables seraient problématiques pour les régressions utilisées par la suite. L'objectif est aussi de sélectionner des variables qui représentent les axes explicatifs de l'ACP, notamment en fonction des coefficients d'explication du jeu de données par les axes. En fonction de ce coefficient, on se focalise sur un certain nombre de dimensions de l'ACP. Enfin, il est également possible de sélectionner des variables ayant un sens écologique au regard de l'objet étudié. Dans cette étude, on s'intéresse à un réseau de dispositifs expérimentaux, et non à une espèce vivante, ainsi, ce critère de sélection de variable ne s'applique pas ici.

Les variables sélectionnées à l'issue de l'ACP sont donc :

- Bio1 : « mean annual temperature »
- Bio2 : « mean diurnal range » (moyenne mensuelle des valeurs extrêmes de température)
- Bio5 : « max temperature of warmest month »
- Bio12 : « mean annual precipitation »
- Bio15 : « precipitation seasonality » (coefficient de variation)
- Bio18 : « precipitation of warmest quarter »

Ce jeu de variables inclut donc 3 variables liées aux variations de température, ainsi que 3 variables liées aux variations de précipitation.

Ensuite, afin de tester si des phénomènes de multi colinéarité existent entre ces 6 variables, la fonction *vif* du package *usdm* a été utilisée. On obtient ainsi le facteur d'inflation

de la variance et la fonction indique si des problèmes de colinéarité existent entre les variables sélectionnées.

6.3. Modèles utilisés

La modélisation des aires de compatibilité climatiques présentes et futures a été effectuée grâce au package *biomod2*, implémenté dans R. BIOMOD est une plateforme informatique qui permet, à l'aide de plusieurs SDMs, de prédire la distribution d'une ou de plusieurs espèces, tout en permettant d'avoir un aperçu des différentes sources d'incertitudes issues des modèles (Thuiller et al., 2009). Cette plateforme permet notamment de créer des modèles d'ensemble qui prennent en compte les résultats de plusieurs algorithmes de distribution selon leurs performances. Cette fonctionnalité est intéressante notamment au regard de la grande variabilité qui peut exister entre les résultats des différents algorithmes.

Pour ce travail, 5 algorithmes différents ont été utilisés : GLM, GBM, GAM, RF et Maxent (nous ne développerons pas ici le mécanisme précis des modèles). La sélection des modèles s'est faite selon leur mode de fonctionnement, en essayant de varier les types de modèles utilisés, tout en se référant à leur importance dans la littérature scientifique. On a utilisé ici des algorithmes dits de « présence/absence », c'est-à-dire que des points de présence et des points d'absence sont en entrée du modèle (contrairement à des algorithmes de « présence unique ») (Pecchi et al., 2019).

Deux des algorithmes utilisés sont basés sur la régression statistique. Les Generalized Linear Models (GLM) sont largement utilisés et répandus en écologie, car très flexibles et simples d'utilisation avec des données aux distributions variées (Guisan et al., 2002). Les Generalized Additive Models (GAM) sont aussi largement utilisés (Pecchi et al., 2019), et présentent une plus grande flexibilité encore que les GLM, notamment lorsqu'il n'existe pas de combinaisons linéaires entre les variables d'entrée.

Deux autres algorithmes sont basés sur l'apprentissage automatique (machine learning). Les Generalized Boosted Models (GBM) ou Boosted Regression Trees (BRT), très utilisés dans le domaine de l'écologie (J. Elith et al., 2008), utilisent une méthode de sélection des données nécessaires à la construction des arbres de décision appelée « boosting method ». L'algorithme Random Forest (RF) est l'un des plus utilisés, et pour cause, c'est l'un des plus performants (Pecchi et al., 2019). La méthode de construction des arbres de décision est aléatoire. Le paramètre intéressant de cet algorithme est qu'il est non vulnérable à des phénomènes de colinéarité entre variables.

Enfin, l'algorithme MaxEnt, un peu différent des autres dans son fonctionnement car décrit comme un algorithme de « presence only », ne nécessitant donc pas de données d'absences (Jane Elith et al., 2011). Pecchi et al. (2019) classent cependant ce modèle dans la catégorie des « presence-absence » car il fonctionne avec des « background data », ou « pseudo-absences ». L'algorithme se base sur le principe du maximum d'entropie (non détaillé ici). L'intérêt pour cet algorithme notamment dans le domaine de la recherche en foresterie a largement augmenté ces 10 dernières années.

6.4. Démarche de modélisation

Le paramétrage des modèles a largement été inspiré de l'ouvrage « Habitat Suitability and Distribution Models : with applications in R » écrit par Guisan, Thuiller et Zimmerman (2017). Trois paramétrages ont été effectués, en fonction des jeux de données utilisés (les jeux de données GRECO et ESP ont eu le même paramétrage, car similaires dans le nombre d'individus).

i) Formatage

Le formatage est une étape importante car c'est à ce moment-là que l'on décide de l'utilisation des pseudo-absences. Pour cette étude, le choix a été effectué après avoir dans un premier temps modélisé des aires de compatibilité au moyen de couches vectorielles, donc en créant des points d'absence manuellement. Le biais dans cette méthode repose dans le fait que les points d'absence étaient répartis sur le territoire français selon une couche de points séparés de 10 km chacun, formant un réseau aligné. Les points étaient répartis systématiquement, et non aléatoirement, ce qui a poussé à utiliser la méthode des pseudo-absences générées aléatoirement selon l'emprise des rasters.

Les pseudo absence sont donc des points d'absence choisis de manière aléatoire sur l'emprise sélectionnée. Cela permet d'utiliser des algorithmes type « absence/présence » tout en évitant les biais qui existent lors de la collecte et l'échantillonnage de données de « vraies absence » (Barbet-Massin et al., 2012). En fonction de l'effectif des données de présence, des algorithmes utilisés, le nombre de pseudo-absence ainsi que leur méthode de génération peut avoir un impact sur la performance des modèles. Ainsi, le paramétrage a été quelque peu différent en fonction du jeu de données sélectionné. En effet, du fait de la taille du jeu de données FRANCE, 10000 pseudo-absence ont été privilégiées (10 fois plus de pseudo-absences que de présences). Pour les jeux de données GRECO et ESP, 4000 pseudo-absence ont été privilégiées (plus de 10 fois le nombre de présence).

Les paramètres de sélection ont des performances différentes en fonction des modèles utilisés (régression vs machine learning). Des compromis ont dû être fait sur ces paramètres notamment car les algorithmes doivent tourner ensemble dans le package *biomod2* afin de pouvoir créer des modèles d'ensemble par la suite. Ainsi, la méthode de sélection 'random' (plus efficace avec les GLM, GAM) a été privilégiée à 'SRE' (plus efficace avec RF, GBM). Un nombre élevé de pseudo-absence est par ailleurs plus efficace avec des algorithmes de régression, tandis qu'un nombre égal au nombre de présences est préférable avec des algorithmes de machine learning. Enfin, 2 sets de pseudo-absence ont été générés pour assurer une répétabilité.

ii) Construction de modèles individuels

Les paramètres de modélisation sont les mêmes sur les 3 jeux de données. Les points-clés de cette étape sont les modèles utilisés, la séparation des données (*datasplit*), le nombre de runs effectués, ainsi que la méthode d'évaluation des performances. Le paramètre *datasplit* permet de séparer le jeu de données selon un certain pourcentage, qui sera utilisé pour calibrer le modèle, le reste des données servant à évaluer le modèle. Ici, un *datasplit* = 70% a été choisi sur les conseils de Wilfried Thuiller. Pour chaque jeu de données, 3 runs ont été effectués. Enfin, les méthodes d'évaluation sélectionnées sont 'ROC' et 'TSS', pour leur prépondérance dans la littérature scientifique. Ces indicateurs se situent entre 0 et 1, plus la valeur est proche

de 1, plus la prédiction est qualifiée de bonne. A l'inverse, un indicateur proche de 0 reflète une mauvaise prédiction. Pour résumer, 30 modèles sont donc construits : 5 algorithmes x 2 sets de pseudo-absences x 3 runs.

iii) Construction de modèles d'ensemble

Une des fonctionnalités intéressantes de *biomod2* est la possibilité de créer des modèles d'ensemble. Ces modèles sont en quelque sorte des combinaisons des modèles individuels construits précédemment, selon leurs performances respectives. Cela permet en partie de tamponner la variabilité qui existe entre les résultats des différents modèles. Il existe plusieurs méthodes pour combiner les modèles. Pour ce travail les méthodes appelées « Models committee averaging » et « weighted mean of probabilities » ont été utilisées, car elles sont communément utilisés dans les études de distribution d'espèces, avec BIOMOD notamment (Hao et al., 2019). Dans cette étape, il est possible de choisir une métrique d'évaluation des modèles, qui permettra de choisir quels modèles vont être conservés pour les modèles d'ensemble. Ici, on a choisi 'TSS' car il semble que cet indicateur soit plus robuste que 'ROC' (ALLOUCHE et al., 2006). Par ailleurs, on décide d'un seuil en-deçà duquel on élimine les modèles construits. Pour le jeu de données FRANCE, on ne détermine pas de seuil, pour les jeux de données GRECO, le seuil TSS est de et ESP, le seuil TSS choisi est de 0.8 (on élimine les modèles dont le score TSS est inférieur à 0.8).

iv) Projections

Après avoir réalisé les modèles d'ensemble, on réalise tout d'abord une projection spatiale des modèles produits, sous scénario actuel. Ensuite, on introduit pour chaque jeu de données les variables environnementales futures sous forme de rasters, pour réaliser des projections sous scénarii futurs (RCP 2.6 et RCP 8.5). On obtient des cartes de compatibilité climatique du réseau GEN4X continues, avec des probabilités de présence comprises entre 0 et 1, ainsi que des projections binaires (présence-absence) basées sur le seuil qui optimise le score TSS des modèles. Ces projections binaires sont exportées sous forme de rasters à l'échelle 1 pixel = 1km x 1km. BIOMOD permet également d'effectuer des cartes appelées « Species Range Change », qui calculent la différence entre la distribution actuelle et future. Ces cartes sont également exportées sous forme de rasters.

7. Comparaison avec le réseau GEN4X

Les cartes de projection binaire « absence/présence » ont ensuite été comparées aux localisations des dispositifs sur QGIS. Pour chaque jeu de données, les dispositifs ont été catégorisés comme « présents » dans l'aire de compatibilité climatique ou « absent », avec les cartes actuelles, futures sous scénario RCP 2.6 et RCP 8.5. Cette catégorisation a pu être effectuée au moyen de la fonction « *Prélever des valeurs rasters* ». L'objectif de cette étape est de dresser une typologie des dispositifs, selon 2 angles différents :

- Les dispositifs qui se trouvent actuellement dans une zone dont le climat futur est déjà exploré par le réseau GEN4X. Ce sont ces dispositifs qui s'avèrent intéressants à étudier maintenant car ils représentent et explorent déjà aujourd'hui un climat que l'on retrouvera dans le futur sur le territoire français. Ainsi, étudier la croissance et la reproduction dans ces jardins communs est un bon moyen de prévoir les effets du changement climatique sur les essences présentes dans ces dispositifs.

- Les dispositifs qui se trouvent actuellement dans une zone dont le climat futur n'est pas exploré par le réseau GEN4X. Ces dispositifs représentent un enjeu car, grâce à la modélisation, on sait qu'ils se trouveront plus tard dans des zones non représentées par le réseau et dont le climat n'a pas encore été exploré. Cibler ces dispositifs permet notamment d'adapter leur gestion en sachant qu'ils seront peut-être plus vulnérables à l'avenir car exposés à des climats encore non étudiés sur le territoire français et donc potentiellement défavorables.

C'est grâce à cette typologie et notamment au comptage de ces types de dispositifs que la représentativité climatique du réseau GEN4X a été évaluée. L'objectif est également de cibler les zones climatiques du territoire français qui ne sont pas représentées par des dispositifs expérimentaux, dans une perspective d'installation de nouveaux jardins communs et de gestion forestière.

Résultats

1. Analyse de la base de données GEN4X

Cette description concerne la base de données complète et non les jeux de données évoqués précédemment.

1.1. Description de la base

Tout d'abord, une analyse temporelle du nombre de plantations permet de mettre en évidence les tendances dans les pratiques de recherche en génétique forestière effectuée par INRAE (figure 6).

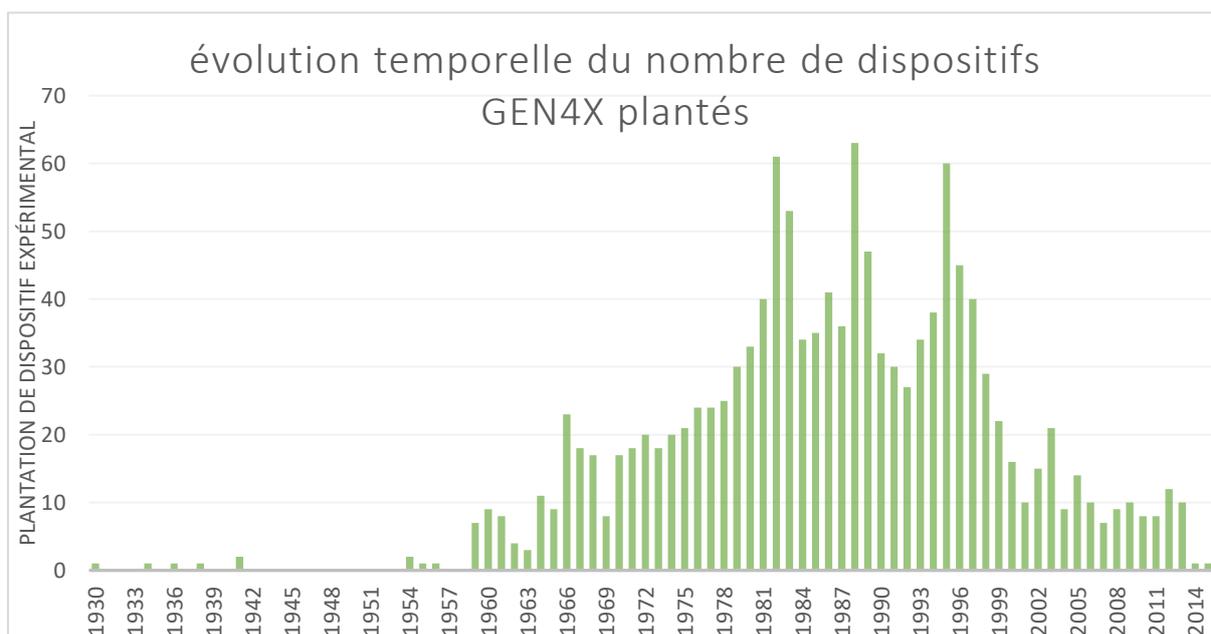


Figure 6 : histogramme du nombre de dispositifs GEN4X plantés au cours du temps

Les premières plantations, datent des années 1930. Une nette augmentation des plantations de jardins communs est observée entre les années 60 et les années 90, pour atteindre un nombre maximum en 1988 avec 63 dispositifs plantés. Parmi ces 63 dispositifs, l'espèce la plus plantée est l'épicéa commun (*Picea abies*), avec 25 plantations, suivie du merisier (*Prunus avium*), avec 10 plantations. A partir de la fin des années 90, le nombre de plantations par an diminue nettement. Bien que la base de données n'aille pas au-delà de 2017, le nombre de jardins communs installés depuis reste très limité.

On compte 92 espèces représentées dans le réseau, ce qui représente au moins 32 genres différents (certains dispositifs ne peuvent être caractérisés car ils comportent la mention « mélange d'espèces ») (figure 7). Les genres les plus représentés sont majoritairement des conifères : mélèzes (*Larix*), pins (*Pinus*), sapins (*Abies*), douglas (*Pseudotsuga*) ou encore épicéas (*Picea*). Le genre de feuillu majoritaire est le peuplier (*Populus*). (Voir Annexe pour la liste complète des espèces).

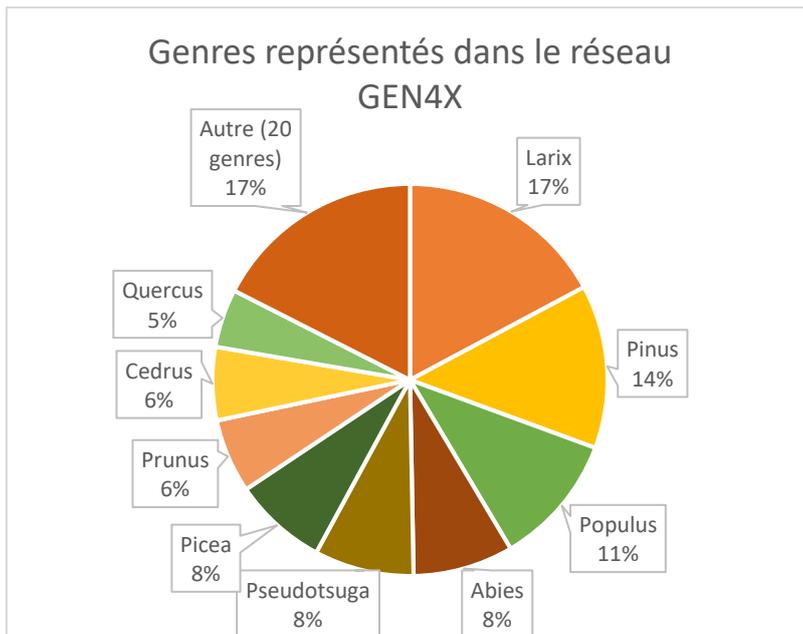


Figure 7 : Proportion de genres représentés dans les dispositifs GEN4X

Concernant les types de dispositifs représentés, on compte un nombre plus élevé de tests de comparaison de descendance (474), que de tests de comparaison de provenances (416). Le réseau comporte une plus faible proportion de tests de comparaison de clones (274), et très peu de tests de comparaison d'espèces (24). (Voir Annexe pour la liste complète).

En ce qui concerne les surfaces des dispositifs expérimentaux, les mesures ont été effectuées sur 1185 individus par défaut d'information. On relève une surface moyenne d'environ 2 ha. La surface maximale est atteinte avec un test de

comparaison de descendance à Walmath (87) avec 30,5 ha. A l'inverse, la surface minimale répertoriée concerne 6 dispositifs et atteint 0.25 ha. La surface médiane est de 1.35 ha.

1.2. Analyse spatiale par sylvoécotérritoires

Sur 86 sylvoécotérritoires référencés sur le territoire français métropolitain, on compte 12 sylvoécotérritoires non représentés par des dispositifs (voir figure 8).

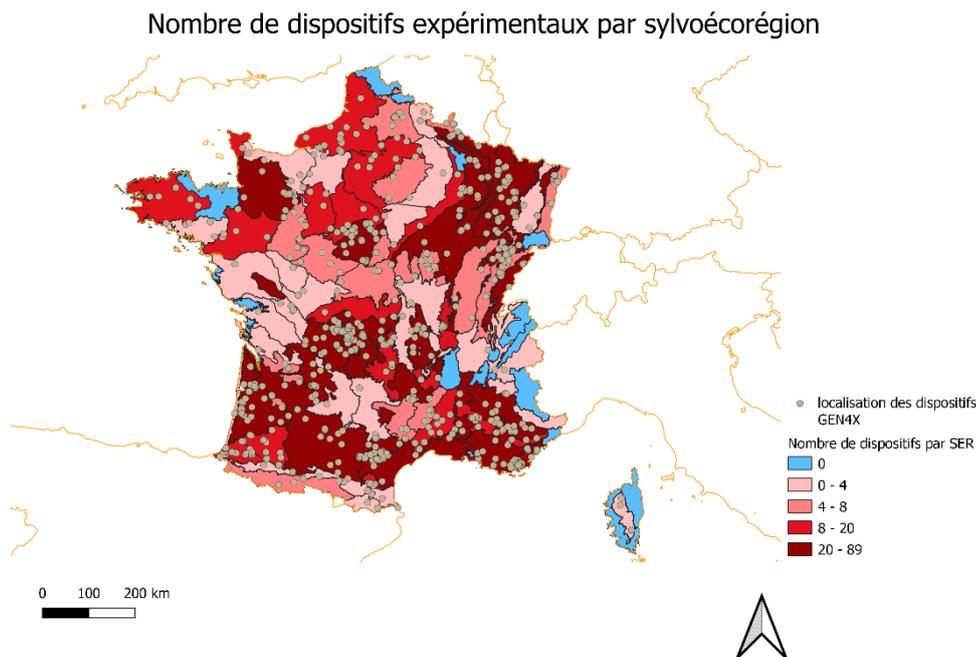


Figure 8 : Nombre de dispositifs expérimentaux GEN4X par sylvoécotérritoires

Les sylvoécotérritoires non représentés par des dispositifs (en bleu sur la carte) sont localisés sur l'ensemble du territoire français et ne semblent pas suivre de tendance particulière. Les SER B70 (Sologne-Orléanais) et C30 (Plaines et dépressions argileuses du Grand Est) sont celles qui comptabilisent le

plus de dispositifs, avec 89 chacune.

2. Exploration des données climatiques

Les résultats de l'exploration des données climatiques sont visualisés sur les diagrammes suivants.

2.1. Diagramme température/précipitations

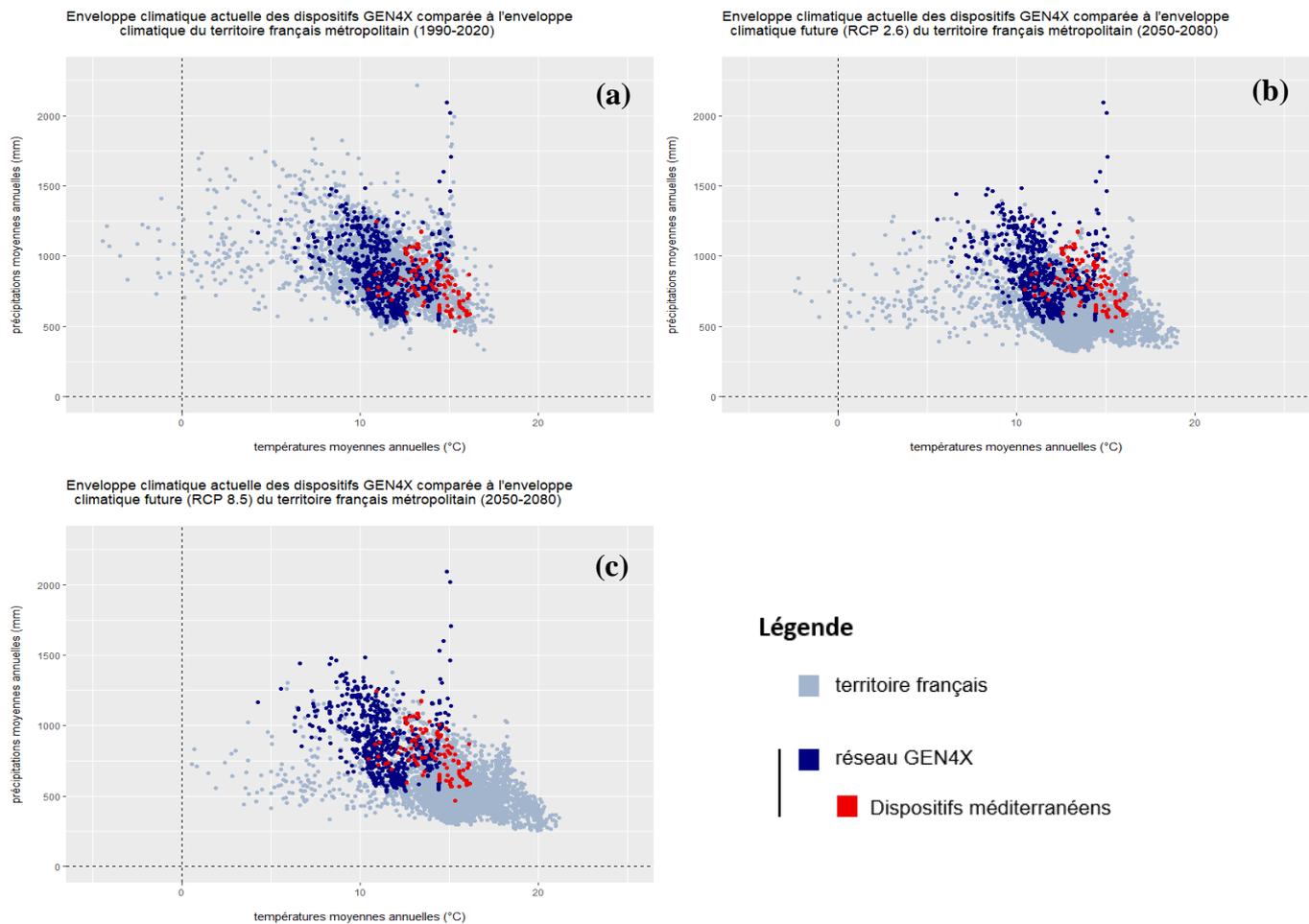


Figure 9 : Diagrammes température/précipitation comparant les climats du territoire français actuels (a) et futurs sous RCP 2.6 (b) et RCP 8.5 (c) avec le climat actuel des dispositifs GEN4X (avec une mise en évidence des dispositifs méditerranéens situés dans les GRECO au climat méditerranéen)

La représentativité climatique du réseau GEN4X est évoquée ici en termes de températures moyennes annuelles et précipitations moyennes annuelles moyennées sur 31 ans (figure 9). Le diagramme (a) suggère une bonne représentativité du réseau GEN4X, avec des dispositifs méditerranéens situés dans les valeurs hautes de température. On remarque un décalage des températures françaises vers des valeurs élevées sur le diagramme (b) ainsi que des précipitations moins élevées, ce qui suggère une perte de représentativité du réseau GEN4X actuel. Le diagramme (c) montre une évolution similaire avec des valeurs de température encore plus élevées (au-delà de 20°C). Les dispositifs méditerranéens actuels seraient donc représentatifs d'un climat français futur moyen, voire plus froid que la moyenne. Par ailleurs, on note que le nuage français (bleu clair) sur les diagrammes (b) et (c) semble plus concentré, ce qui suggère une homogénéisation des températures et précipitations moyennes annuelles.

2.2. Diagrammes de saisonnalité

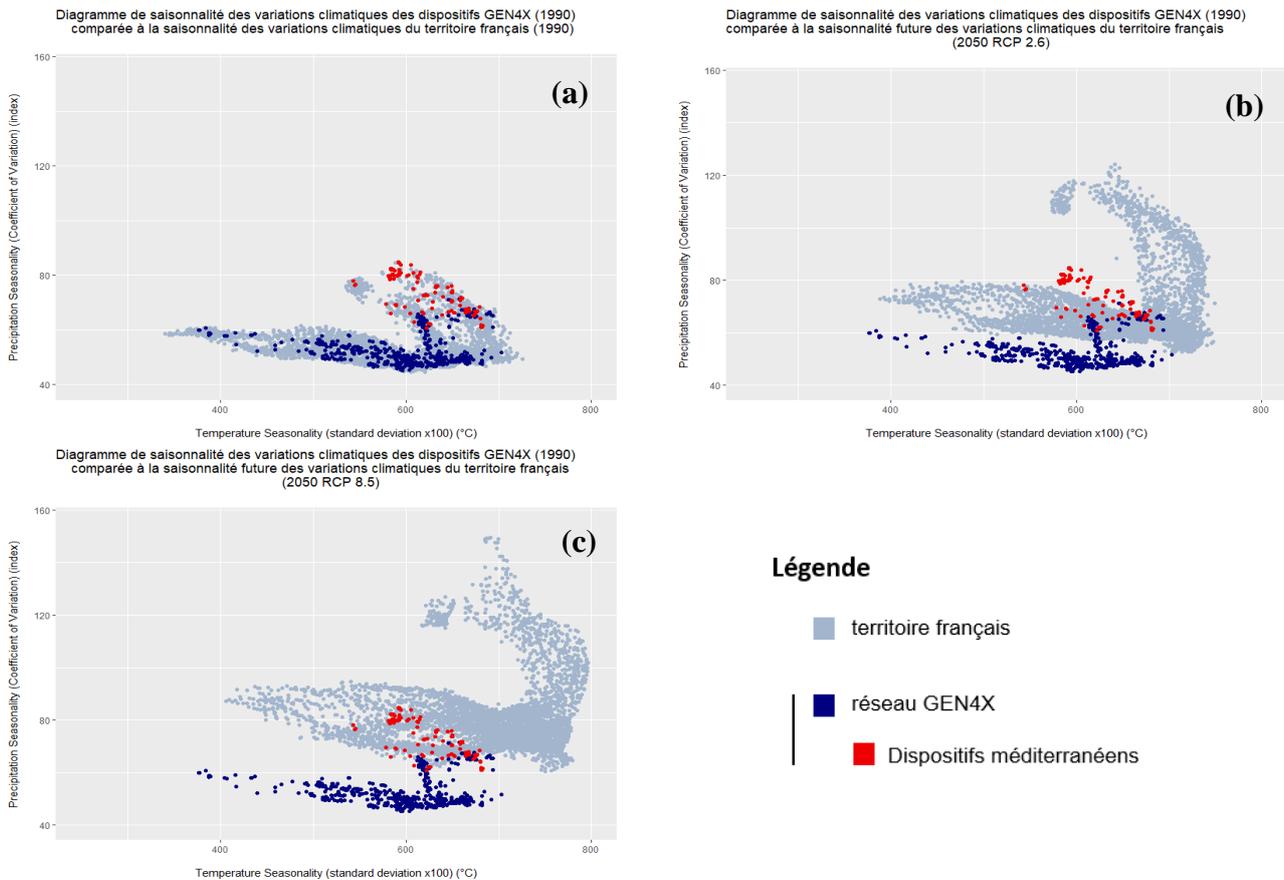


Figure 10 : Diagrammes de saisonnalité comparant les climats du territoire français actuels (a) et futurs sous RCP 2.6 (b) et RCP 8.5 (c) avec le climat actuel des dispositifs GEN4X (avec une mise en évidence des dispositifs méditerranéens situés dans les GRECO au climat méditerranéen)

On observe la représentativité climatique du réseau GEN4X en termes de saisonnalité, c'est à dire de variation entre les saisons (figure 10). De la même manière que pour les diagrammes précédents, on note que le réseau GEN4X couvre bien le nuage français sur le diagramme (a), suggérant une bonne représentation de la saisonnalité actuelle. Néanmoins, le diagramme (b) montre un décalage des valeurs françaises de saisonnalité de précipitation vers des valeurs élevées, ainsi qu'une légère augmentation des valeurs de saisonnalité de température. De la même manière, le diagramme (c) montre une large augmentation de la saisonnalité de précipitation et de températures. Par ailleurs, le nuage français semble de moins en moins concentré, ce qui suggère une saisonnalité moins homogène, et donc des variations de températures et de précipitations plus importantes entre les saisons. Le réseau GEN4X actuel ne semble pas très représentatif des valeurs françaises futures.

2.3. Diagrammes des valeurs extrêmes

Diagramme des valeurs climatiques extrêmes des dispositifs GEN4X (1990)
comparée aux valeurs climatiques extrêmes du territoire français (1990)

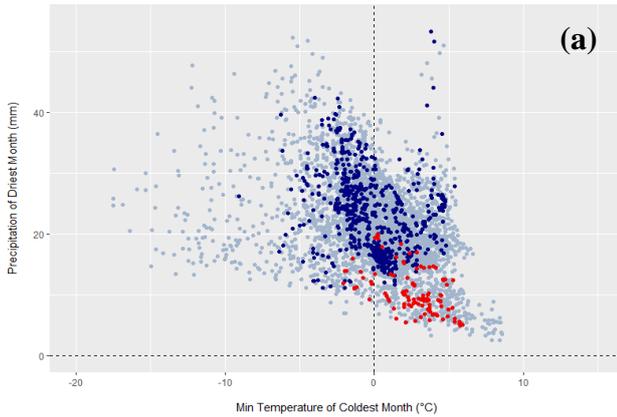


Diagramme des valeurs climatiques extrêmes des dispositifs GEN4X (1990)
comparée aux valeurs climatiques extrêmes futures du territoire français (2050 RCP 2.6)

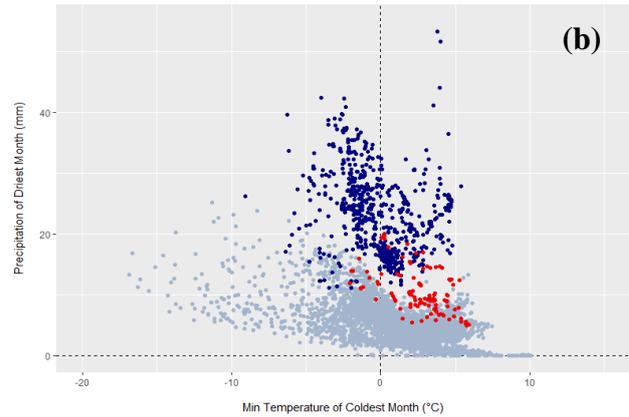
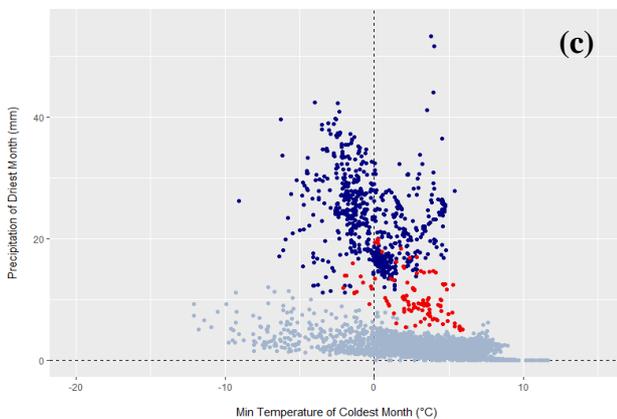


Diagramme des valeurs climatiques extrêmes des dispositifs GEN4X (1990)
comparée aux valeurs climatiques extrêmes futures du territoire français (2050 RCP 8.5)



Légende

- territoire français
- réseau GEN4X
- Dispositifs méditerranéens

Figure 11 : Diagrammes de valeurs minimales de températures du mois le plus froid de l'année et précipitations du mois le plus sec comparant les climats du territoire français actuels (a) et futurs sous RCP 2.6 (b) et RCP 8.5 (c) avec le climat actuel des dispositifs GEN4X (avec une mise en évidence des dispositifs méditerranéens situés dans les GRECO au climat méditerranéen)

On observe ici la représentativité du réseau GEN4X en termes de valeurs de températures et de précipitations extrêmes (figure 11). Ainsi, le diagramme (a) montre une bonne représentation de ces valeurs par le réseau sous climat français actuel. Sous climat futur, cette représentativité diminue et devient quasi-nulle. En effet, le diagramme (b) montre une large diminution des valeurs précipitation lors du mois le plus sec. Le diagramme (c) montre une diminution encore plus importante de ces valeurs de précipitation, ainsi qu'une augmentation des valeurs minimales de température du mois le plus froid. Le réseau GEN4X actuel ne semble ainsi pas représentatif des phénomènes climatiques extrêmes qui risquent de se produire dans un futur proche.

3. Modélisation des aires de compatibilité climatique du réseau GEN4X

Après avoir décrit la base de données de référence pour ce travail, et exploré les données climatiques actuelles et projetées grâce à ClimateDT, une modélisation des aires de compatibilité climatique du réseau GEN4X a été effectuée afin de prédire spatialement la représentativité du réseau quant aux climats français actuels et futurs.

3.1. Sélection des variables d'intérêt

Parmi les 19 variables climatiques projetées grâce à ClimateDT, une sélection a été effectuée au moyen d'une Analyse en composantes principales. L'ACP a donc été réalisée sur 19 variables quantitatives actives.

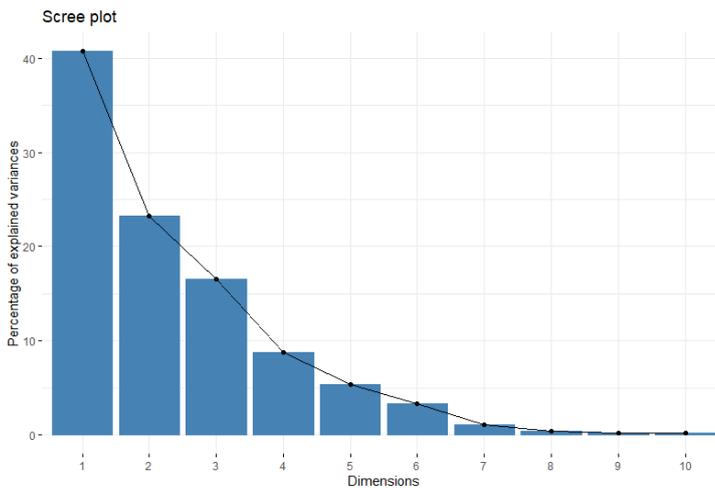


Figure 12 : Pourcentage de variance expliquée par les 10 premières dimensions de l'ACP

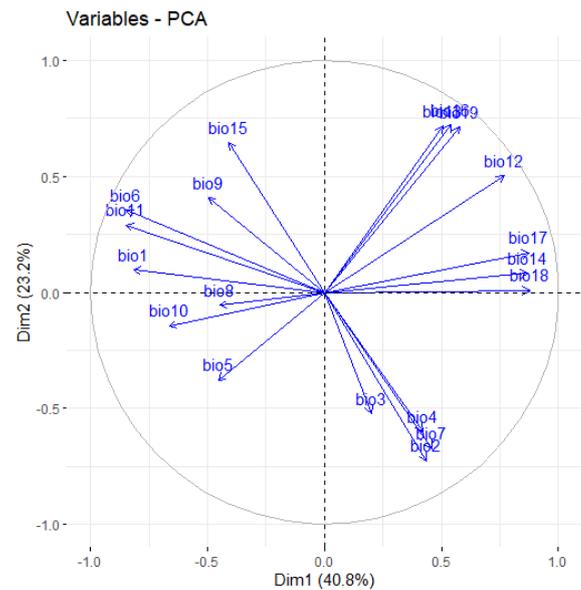


Figure 13 : Cercle des corrélations représentant la contribution des variables climatiques par rapport aux dimensions 1 et 2 de l'ACP

Les deux premières dimensions de l'ACP expliquent à elle deux 64% de l'inertie totale du jeu de données (figure 12). On en déduit que 64% de la variabilité totale du jeu de données constitué par les pixels des 19 variables climatiques est représenté dans le plan de dimensions 1 et 2. Ce total est élevé, on peut ne conserver que ces deux dimensions pour expliquer les liens qui existent entre les variables. Il a été décidé de sélectionner des variables en fonction de leur position relative sur le cercle et de leur distance à l'origine (figure 13). Les variables proches du cercle (dont le \cos^2 est proche de 1) représentent mieux les composantes. Ainsi, comme indiqué dans la partie Matériel et Méthodes, les variables sélectionnées pour le travail de modélisation sont : bio1, bio2, bio5, bio12, bio15 et bio18. Par ailleurs, ces variables ne présentent pas de problèmes de colinéarité.

3.2. Résultats de BIOMOD

3.2.1. Modèles individuels

Les sorties de la fonction de modélisation individuelle par algorithme pour chaque jeu de données sont affichées dans la figure 14. Les performances par algorithmes sont évaluées par les scores ROC et TSS. On note une forte hétérogénéité dans les performances des algorithmes pour le jeu de données France, avec les algorithmes type « machine learning » (RF, GBM) plus performants que les algorithmes de régression (GLM, GAM). Les scores restent peu élevés.

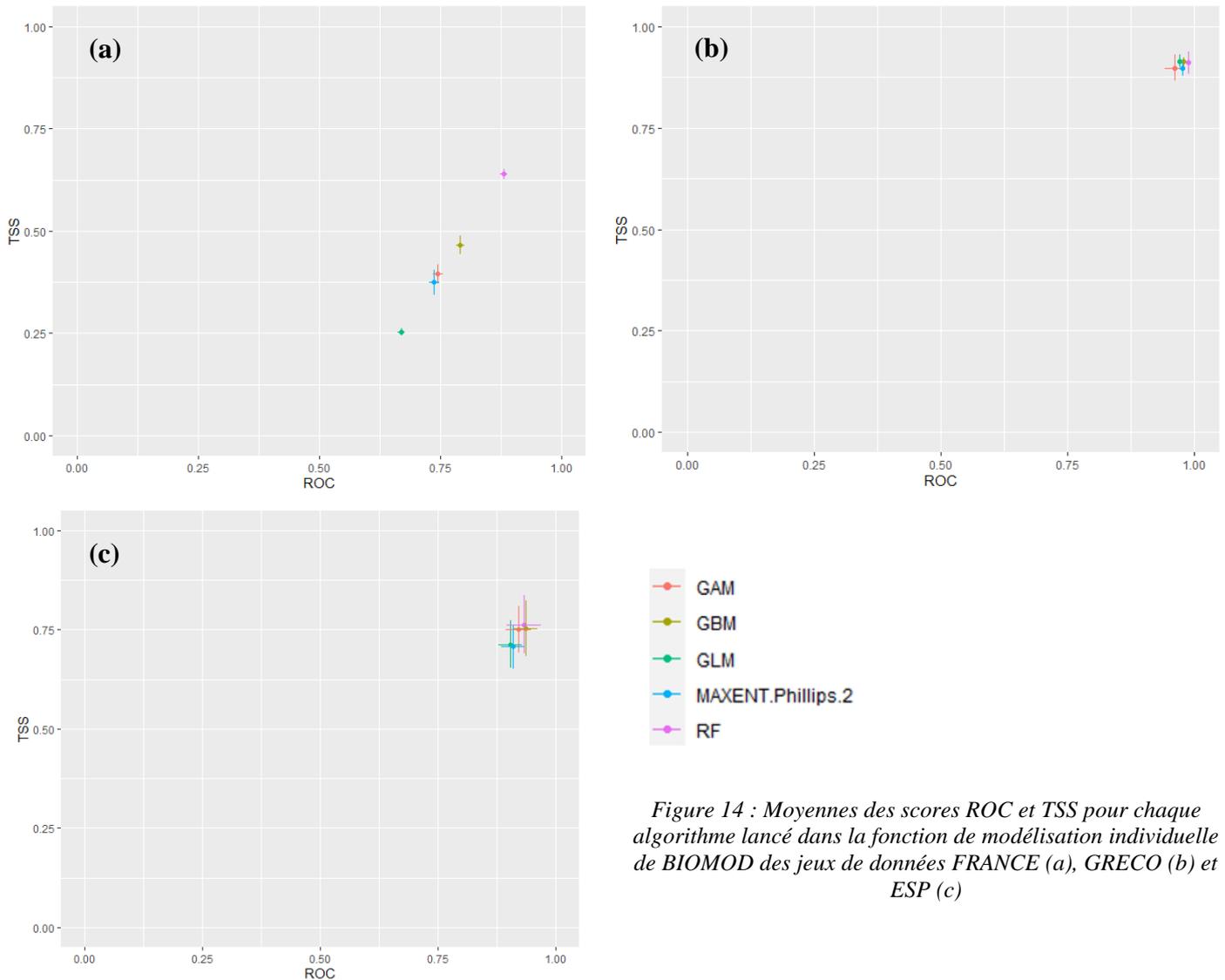


Figure 14 : Moyennes des scores ROC et TSS pour chaque algorithme lancé dans la fonction de modélisation individuelle de BIOMOD des jeux de données FRANCE (a), GRECO (b) et ESP (c)

A l'inverse, les scores des algorithmes pour les jeux de données GRECO et ESP sont très homogènes et répartis dans des valeurs hautes de scores ROC et TSS. La déviation standard (taille des barres d'incertitude) pour le jeu de données ESP est toutefois plus élevée que pour le jeu de données FRANCE. Par ailleurs, les scores ROC et TSS ont également été évalués par sets de pseudo-absence et par runs pour les 3 jeux de données. De la même manière, on note des scores plus bas pour le jeu de données FRANCE. Cependant, par jeu de données, ces scores sont très homogènes entre les différents sets de pseudo-absence

et les différents runs. Cela indique une faible variabilité due à la sélection aléatoire des pseudo-absence ainsi qu'une faible variabilité entre les runs des modèles individuels.

3.2.2. Modèles d'ensemble

Après avoir initié la modélisation avec les modèles individuels, des modèles d'ensemble sont créés afin de tamponner la variabilité existante entre les résultats des différents algorithmes. Le tableau X rassemble les scores TSS et ROC des modèles d'ensemble selon les méthodes « committee averaging » (ca) et « weighted mean » (wm).

	FRANCE		GRECO		ESP	
	ca	wm	ca	wm	ca	wm
ROC	0.867	0.944	0.989	0.994	0.969	0.977
TSS	0.551	0.728	0.942	0.946	0.821	0.846

Tableau 1 : Scores ROC et TSS pour chaque modèle d'ensemble en fonction du jeu de données

Les scores TSS sont sensiblement plus faibles que les scores ROC. Cependant, on se concentre uniquement sur le TSS désormais, qui a été choisi comme métrique d'évaluation des modèles. Ainsi, on note de meilleurs résultats sur les modèles d'ensemble pour les jeux de données GRECO et ESP que pour le jeu de données FRANCE. On remarque notamment que les scores dépassent 0.9 pour le jeu de données GRECO, ce qui témoigne d'une grande efficacité prédictive des modèles d'ensemble pour ce jeu de données. Par ailleurs, les scores sont quelque peu plus élevés pour le modèle d'ensemble weighted mean que pour committee averaging.

3.2.3. Projections

La figure 15 montre les projections actuelles et futures sous scénario RCP 2.6 et 8.5 de la distribution de l'espèce « GEN4X », ou plus exactement des aires de compatibilité climatique des dispositifs expérimentaux, ce qui représente les zones climatiques qui sont ou seront explorées par des dispositifs expérimentaux actuels. Les 3 jeux de données y sont représentés.

Pour le jeu de données FRANCE (les 6 cartes de gauche), les prédictions entre Committee averaging et Weighted mean models semblent peu différentes en termes de zones de distribution avec une forte probabilité de présence. Pour les projections actuelles, la distribution semble cohérente avec les zones de forte densité de présence de dispositifs (zone méditerranéenne, côte aquitaine, massif central, bassin parisien, grand est). Sous scénario 2.6 les zones de la partie nord du territoire s'effacent pour laisser place à des zones de présence sur une grande partie de la côte atlantique. Sous scénario 8.5, la péninsule bretonne est bien représentée, ainsi que la Normandie mais aussi les territoires frontaliers à l'est correspondant à des zones de moyenne et hautes altitudes (Vosges, Jura, Alpes).

Concernant le jeu de données GRECO (les 6 cartes centrales), les prédictions semblent plus pessimistes pour le Weighted mean que pour le Committee averaging. La zone de distribution actuelle correspond aux GRECO J et K (avec une petite zone en Alsace). Les projections futures indiquent une aire de compatibilité qui s'étend sur le territoire français sous RCP 2.6 (massif

FRANCE

GRECO

ESP

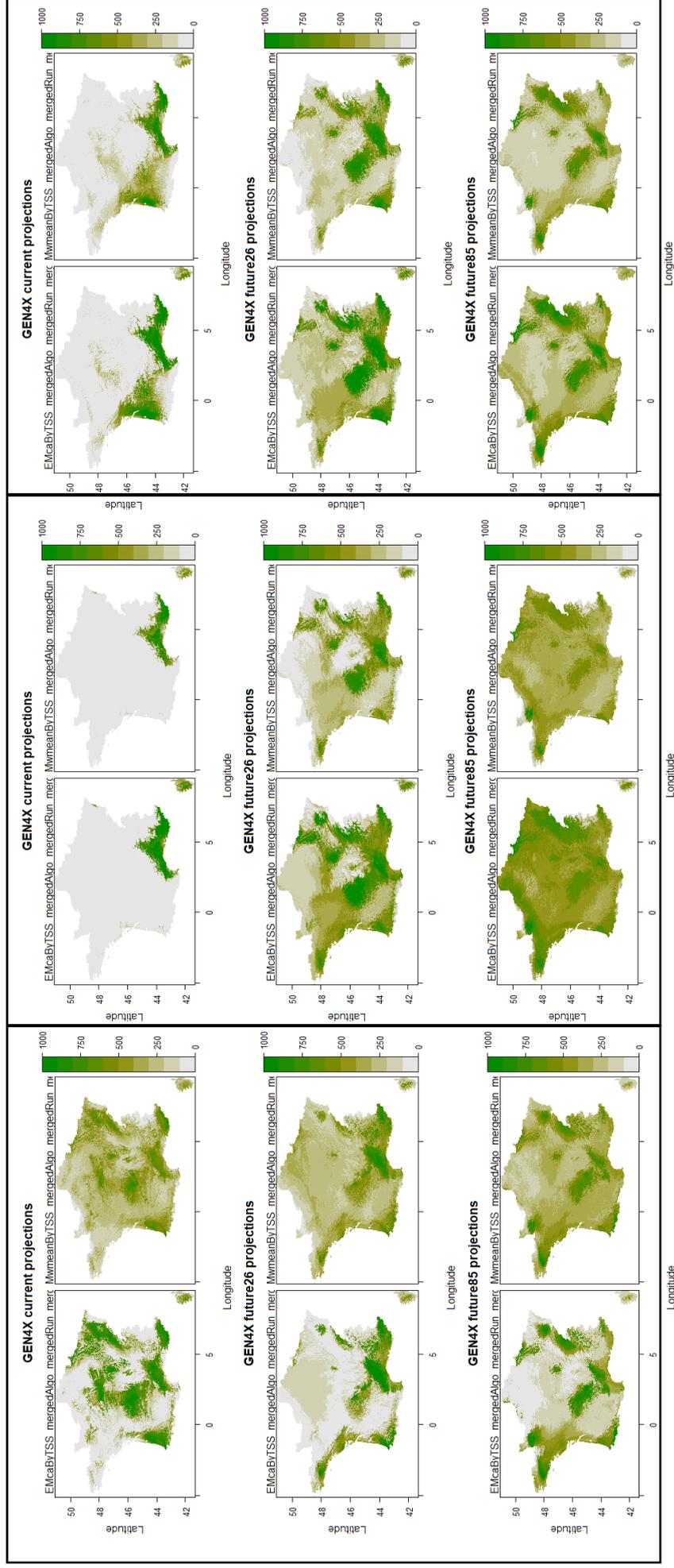


Figure 15 : Cartes des probabilités de compatibilité climatique modélisées en fonction de jeux de données. Les 6 cartes de gauche correspondent au jeu de données France, les 6 cartes du milieu correspondent au jeu de données GRECO et les 6 cartes de droite correspondent au jeu de données ESP. Par jeu de données, on retrouve les 3 périodes étudiées (de haut en bas : actuel, RCP 2.6 et RCP 8.5) On retrouve également les deux méthodes de modélisation d'ensemble (à droite Committee averaging, à gauche Weighted mean)

central, Bretagne, Aquitaine, Est...) et qui semble quitter la zone climatique méditerranéenne. La projection sous RCP8 8.5 prévoit des zones de compatibilité situées surtout en bordure du territoire (côte atlantique, frontière est...), mais qui semble quitter les territoires centraux. Ainsi, l'aire de comptabilité climatique des dispositifs situés en GRECO méditerranéenne apparaît comme bien différente de l'aire de compatibilité du réseau entier (jeu de données FRANCE), en fonction des scénarii considérés.

Enfin, pour le jeu de données ESP (les 6 cartes de droite), les prédictions semblent très similaires entre Weighted mean et Committee averaging. La zone de distribution actuelle correspond à celle du jeu de données GRECO mais avec en plus une majeure partie du territoire aquitain. Les prédictions sous RCP 2.6 annoncent également un décalage des zones de compatibilité climatique vers le nord du territoire : le massif central, la frontière montagneuse à l'est ainsi que la péninsule bretonne sont représentés. A l'inverse, la zone climatique méditerranéenne ainsi que le bassin aquitain sont moins représentés. Sous RCP 8.5, ces deux dernières zones sont encore moins représentées, et laissent place à la péninsule bretonne ainsi qu'une partie du Cotentin. La frontière Est est quant à elle, de plus en plus représentée.

Les proportions des surfaces de compatibilité climatique par jeu de données et par époque sont consignées dans le tableau 2. Cela représente la surface du territoire français dont le climat est actuellement exploré par les dispositifs en jardins communs.

	FRANCE			GRECO			ESP		
	Actuel	RCP 2.6	RCP 8.5	Actuel	RCP 2.6	RCP 8.5	Actuel	RCP 2.6	RCP 8.5
Proportion de compatibilité climatique (%)	15	13	14	5	16	18	13	24	23

Tableau 2 : Proportion de la surface du territoire français représentant l'aire de compatibilité climatique en fonction du jeu de données et de la période considérée. Ces proportions ont été calculées avec les nombres de pixels « présence » et « absence » des projections binaires « Weighted mean ».

3.2.4. Evolution de la distribution (Species Range Change)

Les cartes de Species Range Change (SRC) intègrent une composante temporelle puisqu'elles indiquent les différences entre les projections actuelles et les projections futures sous RCP 2.6 et 8.5. Elles permettent de cibler sur le territoire français les zones climatiques qui, entre aujourd'hui et la période future considérée :

- 1) ne vont plus être représentées par des dispositifs de GEN4X (« lost »)
- 2) seront encore représentées par des dispositifs GEN4X (« pres »)
- 3) ne seront toujours pas représentées par des dispositifs GEN4X (« abs »)
- 4) seront représentées dans le futur par des dispositifs GEN4X (« gain »)

Les prédictions de changement de distribution varient selon le modèle d'ensemble. C'est pour cela que l'on choisit de s'intéresser au Weighted mean qui possède des scores de prédiction TSS plus élevés que ceux du Committee averaging. Ainsi, comme décrit dans la partie précédente, les zones climatiques favorables aux dispositifs varient entre la période actuelle et les périodes futures. Pour le jeu de données FRANCE, il est difficile de décrire une tendance spatiale de changement de distribution et donc de représentation. Pour les jeux de données

GRECO et ESP, on note tout de même sous les scénarios RCP 2.6 et 8.5 une perte de représentativité des climats de la zone méditerranéenne. Le tableau 3 permet de quantifier ces changements de distribution et donc de représentativité climatique des jardins communs.

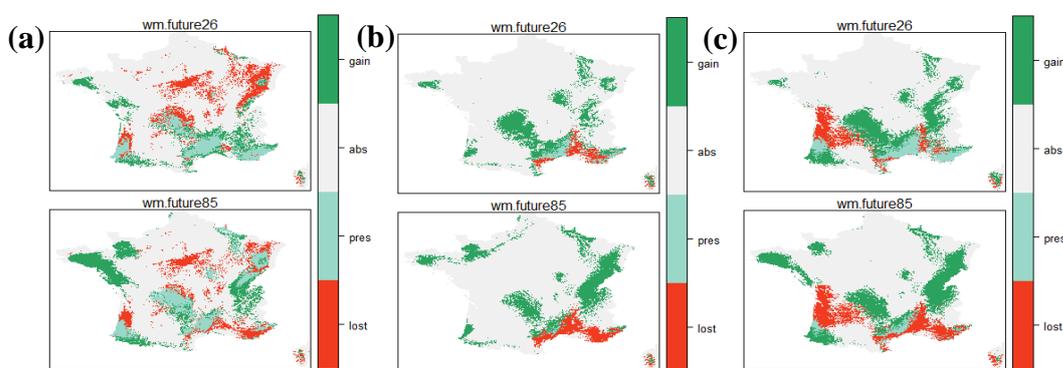


Figure 16 : Cartes de « Species Range Change » qui prédisent l'évolution de la représentativité climatique des jeux de données FRANCE (a), GRECO (b) et ESP (c) entre les périodes actuelles et futures sous RCP 2.6 (cartes supérieures) et entre les périodes actuelles et futures sous RCP 8.5 (cartes inférieures).

Jeu de données	RCP	Lost (pixels)	Abs (pixels)	Pres (pixels)	Gain (pixels)	PercLoss (%)	PercGain (%)	SRC
FRANCE	2.6	90828	720288	65577	56321	58.072	36.010	-22.063
	8.5	87596	706859	68809	69750	56.006	44.596	-11.410
GRECO	2.6	25343	756817	24068	126786	51.290	256.595	205.304
	8.5	44859	721628	4552	161975	90.787	327.812	237.024
ESP	2.6	60586	651282	63281	157865	48.912	127.447	78.535
	8.5	99828	619213	24039	189934	80.593	153.337	72.744

Tableau 3 : Statistiques associées aux cartes de changement de distribution (« Species Range Change »). Pour chaque jeu de données, et chaque RCP, on trouve le nombre de pixels perdus entre la période actuelle et la période future considérée (lost), le nombre de pixels d'absence dont le statut ne change pas (abs), le nombre de pixels de présence dont le statut ne change pas (pres), et le nombre de pixels gagnés entre la période actuelle et la période future considérée (gain). PercLoss représente le pourcentage de la surface actuellement occupée qui est prédit comme étant perdue dans le futur. PercGain représente le pourcentage de la surface occupée dans le futur par rapport à la surface occupée actuelle. SRC représente la différence entre PercGain et PercLoss. Une valeur de SRC négative indique une perte nette de surface de distribution. A l'inverse, une valeur de SRC positive indique un gain net de surface de distribution.

Le jeu de données France est caractérisé par un SRC négatif pour les deux RCP. Cela signifie que la surface de l'aire de compatibilité climatique du réseau GEN4X est susceptible de diminuer dans le futur, et ce sous les deux scénarios RCP. Par ailleurs, il semble que cette surface diminue davantage sous RCP 2.6 que sous RCP 8.5. La surface perdue est en effet comparable entre ces deux scénarios (58% pour le 2.6 et 56% pour le 8.5). En revanche, sous RCP 8.5, la surface gagnée est nettement supérieure (44% contre 36%). Concernant le jeu de données GRECO, on note une large augmentation de sa surface de compatibilité climatique. Sous RCP 8.5, cette augmentation est d'autant plus marquée et l'amplitude entre surface perdue et surface gagnée est d'autant plus élevée. Enfin, le jeu de données ESP révèle une autre

tendance. En effet, tout comme pour GRECO, la prédiction quant au changement de surface de compatibilité climatique est positif, ce qui indique une augmentation de cette surface dans le futur. En revanche, c'est sous RCP 2.6 que ce changement est le plus marqué.

Pour résumer, la représentativité climatique du réseau GEN4X va diminuer entre la période actuelle et la période future et ce, quel que soit le scénario RCP considéré. Cela signifie que l'aire de compatibilité climatique totale représentée et explorée par des dispositifs en jardins communs sur le territoire français est amenée à diminuer. Du point de vue des dispositifs situés dans la région climatique méditerranéenne, c'est l'inverse qui va se produire. En effet l'aire de compatibilité climatique représentée et explorée par des jardins communs méditerranéens est amenée à grandir et à migrer sur le territoire. De la même manière pour les dispositifs en jardins communs qui contiennent des espèces méditerranéennes, cette aire de compatibilité climatique représentée est également amenée à s'étendre sur le territoire français. En revanche, les aires représentées dans le futur seront différentes géographiquement et notamment pour les jeux de données GRECO et ESP, une large perte de représentativité est à prévoir en région méditerranéenne.

3.3. Représentativité climatique du réseau GEN4X

Comme nous l'avons vu précédemment, des changements importants vont avoir lieu dans les aires de comptabilité climatique des dispositifs en jardin communs du réseau GEN4X. Après avoir raisonné en termes de surface, il est intéressant de se pencher sur les localisations, en comparant les aires de compatibilité climatique modélisées précédemment avec les localisations des dispositifs sur le territoire français (tableau 4).

Jeu de données	Période	Nombre de dispositifs présents	Pourcentage de représentativité (%)	Nombre total de dispositifs
FRANCE	<i>Actuel</i>	1067	88,3	1208
	<i>RCP 2.6</i>	306	25,3	
	<i>RCP 8.5</i>	301	24,9	
GRECO	<i>Actuel</i>	179	98,8	181
	<i>RCP 2.6</i>	87	48,1	
	<i>RCP 8.5</i>	18	9,9	
ESP	<i>Actuel</i>	178	96,2	185
	<i>RCP 2.6</i>	113	61,1	
	<i>RCP 8.5</i>	55	29,7	

Tableau 4 : Nombre de dispositifs en jardins communs du réseau GEN4X présents dans l'aire de compatibilité climatique modélisée en fonction du jeu de données et de la période considérée. Le pourcentage de représentativité indique la proportion des jardins communs qui représentent le climat futur. Le nombre total de dispositifs du jeu de données est donné à titre indicatif.

Le jeu de données FRANCE montre une large perte de dispositifs présents en zone de compatibilité climatique future sous scénario RCP 2.6. Cela signifie que 25% des dispositifs du réseau GEN4X actuel seront dans le futur sous RCP 2.6 dans une zone où le climat est actuellement exploré par des dispositifs. La projection sous RCP 8.5 donne un pourcentage de représentativité similaire. Le jeu de données GRECO montre une perte de représentativité sous RCP 2.6 d'environ 50%. Sous RCP 8.5, seulement 10% des dispositifs actuels en jardins communs méditerranéens seront représentatifs du climat actuellement exploré par la partie méditerranéenne du réseau GEN4X. De la même manière, le jeu de données ESP montre une perte de représentativité de 40% sous RCP 2.6. Sous RCP 8.5, c'est seulement 30% des dispositifs actuels étudiant des espèces méditerranéennes qui seront représentatifs du climat

actuellement exploré par la partie du réseau GEN4X qui étudie des espèces méditerranéennes. Cela signifie que 70% des dispositifs du jeu de données ESP est actuellement dans des zones dont les climats ne sont pas explorés par des dispositifs à espèces méditerranéennes.

Le tableau 5 se base sur le même principe de comptage que le précédent mais cette fois-ci en prenant en compte les changements de distributions trouvés grâce au « Species range change ».

Jeu de données	RCP	Perte		Présence		Absence		Gain	
		Compte	%	Compte	%	Compte	%	Compte	%
FRANCE	2.6	771	63,8	296	24,5	130	10,7	10	0,8
	8.5	777	64,3	290	24,0	129	10,6	11	0,9
GRECO	2.6	92	50,8	87	48,1	1	0,5	0	0
	8.5	161	88,9	18	9,9	1	0,5	0	0
ESP	2.6	65	35,9	113	61,1	6	3,2	0	0
	8.5	123	66,4	55	29,7	6	3,2	0	0

Tableau 5 : Nombre de dispositifs en jardins communs du réseau GEN4X présents dans les 4 types de zones des cartes de « Species range change » en fonction du jeu de données et du scénario climatique. La colonne « % » indique la proportion des dispositifs présents dans la catégorie par rapport au nombre total de dispositifs du jeu de données. La catégorie « perte » représente les dispositifs présents actuellement dans l'aire de compatibilité climatique du jeu de données, mais qui ne s'y trouveront plus dans le futur. La catégorie « absence » correspond aux dispositifs déjà absents d'une aire de compatibilité climatique et qui le restera dans le futur. La catégorie « présence » correspond aux dispositifs qui vont rester dans une aire de compatibilité climatique dans le futur. La catégorie « gain » correspond aux dispositifs actuellement non représentatifs du climat actuel et qui vont le devenir dans le futur.

Le jeu de données FRANCE est caractérisé par une perte importante de dispositifs représentatifs des climats actuels dans le futur. Dans les deux projections climatiques 2.6 et 8.5, une baisse de plus de 60% du nombre de dispositifs représentatifs des climats actuels est prédite. Ainsi, seulement le quart des dispositifs en jardin communs du réseau GEN4X va rester représentatifs dans le futur des climats actuels. Par ailleurs, ce jeu de données montre une certaine homogénéité des résultats en fonction du RCP considéré. Le jeu de données GRECO montre que les dispositifs méditerranéens vont voir leur représentativité en climats actuels diminuer de moitié sous RCP 2.6 et d'environ 90% sous RCP 8.5. Cela reflète un large changement de distribution de l'aire de compatibilité climatique pour les dispositifs méditerranéens dans le futur. En effet, seulement 10% du réseau méditerranéen sera encore représentatif des climats actuels. Le reste se situera dans des zones climatiques non explorées par le réseau méditerranéen. Le jeu de données ESP montre une tendance similaire avec une perte de représentativité plus grande sous RCP 8.5 que sous RCP 2.6.

Discussion

Le réseau GEN4X comporte donc une grande diversité biologique, et une diversité de régions représentées (la forêt française comptant actuellement 194 espèces référencées). Le réseau est néanmoins soumis au réchauffement climatique et reste fragile : après 2050, 75% des dispositifs se situeront en dehors de leur niche climatique actuelle. Il semble en revanche que les dispositifs méditerranéens actuels restent représentatifs de la niche climatique du réseau (surtout sous RCP 2.6), mais aussi que l'aire de compatibilité actuelle des dispositifs méditerranéens couvre une certaine partie des dispositifs GEN4X français. Il existe donc un double enjeu pour le réseau : 1) mesurer aujourd'hui les dispositifs qui se situent dans un environnement climatique qui sera inconnu ; 2) suivre sur le long terme les jardins communs qui vont occuper un climat majoritaire après 2050.

1. Performances des modèles

Avant de discuter des résultats de modélisation, il est intéressant de se pencher sur l'évaluation des modèles en fonction des jeux de données. De nombreuses étapes dans le processus de modélisation sont soumises à de la variabilité.

1.1. Réseau entier (jeu de données FRANCE)

La première étape de modélisation, avec des modèles individuels donne des résultats moyens pour ce jeu de données. En effet, si l'on observe les scores ROC et TSS (figure 14) obtenus en sortie de modélisation, on remarque une grande hétérogénéité en fonction du type d'algorithme utilisé, (TSS de 0.25 pour les GLM, 0.6 pour RF). Ces scores sont par ailleurs peu élevés, ce qui témoigne d'une efficacité moyenne des modèles individuels. La variabilité observée peut être due au choix des pseudo-absence. En effet, pour cette étude il a été préconisé de prendre un grand nombre de pseudo-absence avec un petit nombre de sets différents. L'inverse aurait pu être effectué. Par ailleurs, le mode de sélection des pseudo-absence 'random' a été fait par défaut. Il aurait été intéressant de sélectionner par exemple des pseudo-absence en dehors d'un rayon défini autour des points de présence. En effet, ces points sont localisés par des coordonnées décimales qui peuvent être soumis à des erreurs de collection de données ou de saisies. Ce facteur de variabilité n'a pas été pris en compte ici. Cependant, malgré des scores TSS peu élevés par sets de pseudo-absence, ces scores sont similaires, ce qui montre une faible variabilité des résultats induite par les sets de pseudo-absence. De la même manière, les scores des 3 runs par algorithmes et sets de pseudo-absence sont peu élevés, mais similaires. Il est tout de même important de noter des barres d'écart-type importantes. Un autre facteur de performance des modèles est le « data-split ». Il a été choisi ici sur les conseils d'experts en fonction de leur expérience de modélisation.

La création des modèles d'ensemble a permis d'obtenir des modèles d'une plus grande puissance prédictive, en réduisant le risque de produire de mauvaises prédictions. En effet cette technique, de plus en plus utilisée en écologie et évolution, est susceptible de produire des prédictions plus robustes en faisant face à la variabilité des résultats issus des modèles individuels (Araújo & New, 2007). Les scores TSS pour les modèles d'ensemble sont plus élevés (notamment pour la méthode Weighted mean, où il est satisfaisant). Cependant, la création des modèles d'ensemble a pu être possible sans indiquer de seuil de score TSS comme sélecteur de modèles individuels. Ainsi, tous les modèles individuels ont été pris en compte et pondérés selon leur score TSS pour donner les modèles d'ensemble. Cette étape a pu contribuer à réduire la puissance prédictive des modèles d'ensemble.

Le jeu de données FRANCE montre donc une grande variabilité de performances des algorithmes utilisés. Cette variabilité a pu être compensée en partie par la création de modèles d'ensemble, qui possèdent une puissance prédictive supérieure, mais encore moyenne pour ce genre de modèles. La répartition spatiale des points de présence sur le territoire étudié peut être un facteur de variabilité des résultats. En effet, le jeu de données montre une répartition des points sur l'ensemble du territoire français, avec des zones de plus ou moins faible densité. La plateforme de modélisation BIOMOD est dédiée à la modélisation et la projection de distributions d'espèces. Ici, son usage est détourné pour permettre l'analyse multivariée par des modèles corrélatifs de la représentativité climatique d'un réseau de points qui représentent des dispositifs expérimentaux. Ainsi, la répartition des points de présence ne suit pas de « lois biologiques ou écologiques », auxquelles une espèce serait soumise, mais est bien le résultat de 90 ans d'expérimentation en jardins communs sur le territoire français, reflétant les grandes directions de recherche et de production forestière qui furent celles du XX^{ème} siècle. Par ailleurs, le réseau, en couvrant une grande partie du territoire, couvre également des climats très différents, imitant la répartition d'une espèce très ubiquiste et peut-être plus difficile à modéliser avec seulement 6 variables climatiques. C'est pourquoi on peut supposer que cette répartition particulière des points de présence est le facteur principal qui influe sur la puissance prédictive des modèles. Un autre indice de la difficulté de modéliser de manière efficace la niche climatique d'un tel réseau est la présence de points d'absence et de gain (tableau 5), qui devraient être nuls car cela montre que certains dispositifs du réseau (140 en l'occurrence) ne se trouvent pas dans l'aire de compatibilité climatique actuelle du réseau, ce qui semble aberrant.

1.2. Dispositifs au climat méditerranéen (jeu de données GRECO)

Ce jeu de données montre des résultats nettement supérieurs au précédent. Tout d'abord, les modèles individuels semblent déjà très performants. Les résultats entre les algorithmes sont très homogènes et très élevés (score TSS autour de 0.8). Cela montre une prédiction robuste des algorithmes individuels. De la même manière que pour le jeu de données FRANCE, les pseudo-absence ont été choisies en fonction de l'effectif des présences. On a préféré ici prendre des sets importants de pseudo-absences mais avec peu de sets. Ainsi, on remarque que les scores en fonction des sets sont très homogènes, mais aussi élevés. Le choix des sets de pseudo-absence induit donc une faible variabilité dans les résultats. Par ailleurs, le même constat est fait pour les 3 runs différents par algorithmes et sets de pseudo-absence.

Concernant les modèles d'ensemble, on remarque une grande efficacité des prédictions, très équivalentes selon la méthode utilisées (Committee averaging et Weighted mean). La variabilité des algorithmes individuels a pu être compensée, et une sélection a pu être effectuée avec un seuil de 0.8, ce qui est élevé. Les scores TSS sont très élevés (plus de 0.9 à chaque fois), ce qui témoigne d'une prédiction robuste par les modèles d'ensemble. On peut affirmer que les modèles d'ensemble produits à partir du jeu de données GRECO sont plus fiables que ceux produits pour le réseau entier.

Les modèles de distribution de l'aire de compatibilité climatique des dispositifs méditerranéens de GEN4X montrent donc une puissance prédictive importante. Il est donc très probable que les changements de distribution indiqués par les cartes de SRC soient réalisés à l'avenir. Cette prédiction plus fiable que pour le réseau entier vient potentiellement de la répartition initiale des points de présence. En effet, on étudie ici des dispositifs dont la répartition est concentrée dans une zone géographique aux conditions climatiques cohérentes (la zone climatique

méditerranéenne). Cette répartition, en imitant une répartition d'espèce telle qu'on en trouve en milieu naturel, est à l'origine d'une meilleure prédiction de la distribution spatiale et temporelle de la niche climatique des dispositifs méditerranéens.

1.3. Dispositifs aux espèces méditerranéennes (jeu de données ESP)

Le jeu de données ESP est comparable au jeu de données GRECO, en termes de localisation et d'effectif. On remarque que cela influe sur les résultats de modélisation. Les modèles individuels possèdent de bons scores TSS (autour de 0.75), bien qu'un peu inférieurs à ceux trouvés pour le jeu de données GRECO. Les scores sont homogènes, cependant, les valeurs d'écart-type sont les plus élevés pour ce jeu de données. La sélection des pseudo-absence a été la même que pour GRECO, et les résultats sont similaires, indiquant une faible variabilité entre les sets de pseudo-absence et entre les runs.

Les modèles d'ensemble montrent de bons scores (autour de 0.8), bien qu'inférieurs à ceux trouvés pour GRECO. La sélection des modèles individuels à intégrer a été effectuée avec un seuil de TSS de 0.7. On peut donc affirmer que les modèles d'ensemble créés sont robustes et représentent une bonne prédiction de la représentativité climatique des dispositifs à espèce méditerranéenne.

Les performances obtenues avec ce jeu de données viennent confirmer que la puissance prédictive des modèles obtenus est influencée par la répartition des points de présence. En effet, la distribution des points de présence pour ce jeu de données est relativement similaire à celle de GRECO, quoique moins concentrée en région méditerranéenne et plus étalée sur d'autres régions climatiques (forêt des Landes, par exemple). Les modèles d'ensemble bien qu'ayant obtenu des scores TSS très bons sont toutefois un peu inférieurs à ceux obtenus avec le jeu de données GRECO. Par ailleurs, on remarque qu'avec ces deux jeux de données, les dispositifs classés en « absent » ou en « gain » par le Species range change (tableau X), sont très peu nombreux.

2. Représentativité climatique du réseau GEN4X

Cette étude a pour objectif d'évaluer la représentativité climatique du réseau GEN4X c'est-à-dire de déterminer les zones du territoire français dont le climat est exploré par des dispositifs en jardin commun. Par déduction, il est également possible de cibler des zones non-explorées.

2.1. Réseau entier

Le jeu de données FRANCE est le plus important des 3 (1208 individus). Les points de présence de dispositifs sont répartis sur l'ensemble du territoire français et semblent couvrir une grande partie des climats du pays. En effet, seulement 12 sylvoécორégions sur 86 ne sont pas représentées par des dispositifs expérimentaux. Par ailleurs, les diagrammes climatiques réalisés dans les figures 9, 10 et 11 montrent qu'actuellement, le réseau semble couvrir une grande partie de l'enveloppe climatique du territoire français. C'est moins le cas dans les scénarios climatiques futurs, où les dispositifs semblent être de plus en plus écartés de la réalité climatique qui existera après 2050. Les projections climatiques pour ce jeu de données montrent que la surface du territoire français actuellement explorée par des dispositifs en jardins communs ne va pas fondamentalement changer de taille dans le futur. Elle occupe aujourd'hui 15% de la surface du pays, ce qui signifie que 85% de la surface de la France n'est pas explorée. Elle va cependant diminuer quelque peu. En revanche, la distribution de cette aire de

compatibilité sera modifiée : sous les deux RCP, elle va perdre la moitié de sa distribution actuelle et gagner de 35 à 45% de nouvelles surfaces.

Concernant les dispositifs, une large perte de représentativité est à prévoir. En effet elle passe de 88% (idéalement elle devrait être à 100) à 25% environ sous les 2 RCP. Cela signifie qu'après 2050, seulement un quart des dispositifs du réseau GEN4X sera situé dans la niche climatique actuelle du réseau, et donc que 75% fera face à un climat non représenté et exploré par le réseau. Par ailleurs, les 25% de dispositifs encore représentatifs dans le futur semblent être différents en fonction des RCP.

Les résultats de cette partie sont néanmoins à nuancer au regard des performances moyennes des modèles utilisés.

2.2. Partie méditerranéenne du réseau

Le jeu de données GRECO comporte 181 individus (dispositifs). Comme décrit précédemment, les dispositifs sont très concentrés géographiquement et en cela, ce jeu de données intègre une certaine cohérence climatique. Ainsi, on peut considérer que la niche climatique de ces dispositifs représente correctement le climat méditerranéen et son extrapolation spatiale et temporelle.

On trouve donc que les dispositifs situés en région méditerranéenne représentent 5% du climat français total. Cette aire de compatibilité climatique va largement augmenter dans le futur. Ce phénomène de « méditerranéisation » des climats semble cohérent avec ce que l'on trouve dans la littérature (Barredo et al., 2018; Klausmeyer & Shaw, 2009). En effet, après 2050, c'est 16 voire 18% du territoire français dont le climat sera méditerranéen ou en tout cas actuellement exploré par des dispositifs GEN4X de la région. Cela correspond à une augmentation de 250% sous RCP 2.6 et 330% sous RCP 8.5 de la surface explorée sous climat actuel. On retrouve même des résultats similaires à Barredo et al. (2018) en ce qui concerne la « méditerranéisation » de la péninsule bretonne. En revanche, on note que la région méditerranéenne n'est plus représentée dans le futur, ce qui reflète un potentiel réchauffement global et un passage de cette zone géographique à un climat plus aride.

Concernant les dispositifs, sous RCP 2.6 c'est environ la moitié du réseau méditerranéen qui sera encore représentatif du climat actuel de la zone. Sous RCP 8.5, seulement 10% des dispositifs méditerranéens seront encore représentatifs du climat de la région méditerranéenne. Sous ce scénario, cela signifie que 90% des dispositifs GRECO actuels se situeront dans des zones non explorées par le réseau actuel, et probablement aux conditions arides, c'est-à-dire plus chaud et plus sec.

2.3. Dispositifs à espèces méditerranéennes

Contrairement au jeu de données précédent, les dispositifs ne sont pas nécessairement situés en région méditerranéenne. Le climat représenté par cette partie du réseau est difficilement qualifiable car on retrouve autant des dispositifs en région méditerranéenne que dans le bassin aquitain (climat type Sud-Ouest océanique). Aujourd'hui, 13% du territoire français est exploré par des jardins communs à espèces méditerranéennes, en termes climatiques. C'est-à-dire que le climat actuellement représenté par des dispositifs contenant des espèces méditerranéennes représente 13% du territoire français. C'est comparable à la surface actuellement explorée par le réseau GEN4X entier (15%). Etant donné les scores de prédiction des modèles pour le jeu de

données entier (FRANCE), on peut supposer que cette surface est sous-estimée due à une moindre fiabilité de ces modèles. La taille de cette aire de compatibilité climatique va augmenter dans le futur pour atteindre 24% du territoire français (RCP 2.6) ou 23 % (RCP 8.5). Cela représente une augmentation de la surface initiale de l'aire de compatibilité de 78% et 72% respectivement. Si l'on se penche sur les dispositifs, on se rend compte qu'après 2050, sous RCP 2.6 et RCP 8.5, la représentativité diminue. Cela signifie que dans le futur, les dispositifs contenant des espèces méditerranéennes seront de moins en moins représentatifs des climats explorés par ces mêmes dispositifs actuellement. Une grande partie se situera dans des zones dont on ignore le climat, du moins pour cette étude. Cela est particulièrement parlant avec ce jeu de données car il intègre une certaine cohérence fonctionnelle entre les espèces testées. Contrairement aux jeux de données précédents, on peut imaginer que les espèces présentes dans ces dispositifs sont amenées à réagir de manière similaire aux perturbations climatiques qu'elles peuvent subir puisqu'elles sont endémiques d'une région au même climat. Ainsi, trouver que 70% des dispositifs à espèces méditerranéennes se situeront dans des zones dont le climat futur est actuellement non testé représente non seulement un enjeu de gestion, car ces dispositifs sont potentiellement amenés à disparaître, mais aussi un enjeu de recherche, car on ne connaît pas leur aptitude à faire face à ces perturbations.

3. Implications de ces résultats et perspectives de recherche

D'une manière générale, il est difficile de mettre en évidence des tendances concernant le réseau entier. Il intègre des dispositifs répartis aux quatre coins de la France et donc lorsque l'on étudie le réseau de manière globale, les pertes de représentativité quelque part sont compensées par des gains ailleurs. Par ailleurs, nous avons vu que la modélisation s'avère délicate avec un tel jeu de données. C'est pourquoi il est difficile de tirer des conclusions lorsque l'on se penche uniquement sur le réseau GEN4X dans son ensemble. En revanche, les résultats trouvés avec les jeux de donnée plus restreints et plus concentrés sont plus exploitables. C'est pourquoi il est intéressant de partitionner le réseau afin d'obtenir des résultats plus fins.

Mettre en perspective les résultats trouvés avec les jeux de données plus restreints permet de comprendre davantage l'évolution de la représentativité climatique des dispositifs en jardins communs dans les années à venir. En effet, lorsque l'on étudie des jeux de données plus petits et plus concentrés géographiquement, les résultats de modélisation sont plus fiables. Cela permet ensuite par exemple de caractériser des zones climatiques. En effet, le jeu de données GRECO représente plutôt bien le climat méditerranéen. Ainsi, la représentativité climatique des dispositifs en jardins communs issus de cette partie du réseau permet de conclure sur les zones dont le climat sera méditerranéen dans le futur sur le territoire français. Confronter ces zones de compatibilité climatique avec le réseau entier permet de qualifier chaque dispositif en fonction du type de climat qu'il représentera dans le futur (en l'occurrence méditerranéen). Ainsi, on observe que 364 dispositifs du réseau GEN4X se situent dans des aires de compatibilité climatique future des dispositifs méditerranéens. A partir de là, il est possible d'adapter les pratiques de gestion et d'expérimentation de ces jardins communs en conséquence. Une piste de recherche serait donc d'effectuer cette même analyse pour toutes les régions climatiques de France, et de confronter les aires de compatibilité de chaque climat testé avec les localisations des dispositifs du réseau. On suppose que cette analyse ne permettra pas

de couvrir l'intégralité du territoire français, car le changement climatique est à l'origine de contractions et d'expansions de zones climatiques (Klausmeyer & Shaw, 2009). C'est pourquoi des données climatiques issues d'autres pays (autour du bassin méditerranéen par exemple) seraient primordiales pour conduire une telle étude. Cette analyse permettrait également de mettre en évidence de manière plus fine les zones du territoire français dont le climat ne serait pas du tout étudié par des dispositifs dans le futur.

Par ailleurs, s'intéresser aux espèces présentes dans chaque dispositif (comme le jeu de données ESP) permettra de confronter les résultats des expérimentations menées sur ces jardins communs avec le changement de climat qui opérera sur ces zones. Des espèces venant de régions au climat plutôt homogène seront davantage susceptibles de se comporter de manière similaire face au changement climatique. C'est pourquoi il serait intéressant (mais coûteux en temps) d'effectuer ce même travail de modélisation en séparant le réseau GEN4X par espèces. On peut ensuite connaître le spectre de climats représenté par une espèce au sein du réseau et, avec des projections, anticiper le comportement des essences dans les différents jardins communs pour véritablement cibler les dispositifs où il est intéressant de continuer les expérimentations, ou d'arrêter d'étudier telle espèce dans telle zone. Par ailleurs, dans une optique de plantation durable, il serait également possible d'envisager cette analyse sous l'angle génétique. En effet, au regard de la grande variabilité intraspécifique qui existe au sein des essences forestières, il serait opportun d'intégrer les climats couverts par les différentes niches climatiques des provenances d'une même espèce testées dans les jardins communs. En confrontant ces niches climatiques avec le climat français futur, on pourrait cibler des dispositifs susceptibles d'accueillir de nouvelles provenances, dont la niche climatique correspond au climat futur en France.

Les jardins communs méditerranéens ont donc un rôle important à jouer dans les années à venir. Ils représentent 15% du réseau entier mais le climat qu'ils occupent actuellement sera amené à s'étendre sur le territoire français. Il semble intéressant de planter des espèces endémiques de la région méditerranéenne, qui prospéreront dans ces nouvelles zones (le massif central, la péninsule bretonne sous RCP 2.6 voire la Normandie mais aussi la frontière est sous RCP 8.5). Ainsi, ils renseignent aujourd'hui sur un climat qui occupera presque 20% du territoire français. Il pourrait également être intéressant de planter des espèces non-méditerranéennes susceptibles de se retrouver au sein de ces nouvelles aires de compatibilité, dans les jardins communs méditerranéens. Cela donnerait un aperçu de ce qui attend ces espèces à l'horizon 2050. Les dispositifs situés en région méditerranéenne sont également soumis à un changement de climat important, car si leur aire de compatibilité climatique va s'étendre, elle va quitter la région actuelle, laissant place à un climat plus chaud et sec. Ces dispositifs, bien que représentant un large potentiel de reboisement, sont donc les témoins de la vulnérabilité des peuplements méditerranéens face au réchauffement climatique.

Enfin, les territoires d'altitude (zones montagneuses) gagneraient également à être investigués, ils semblent pouvoir accueillir des dispositifs, au regard des prédictions climatiques favorables pour le futur. Utiliser les zones de perte de représentativité permettrait également de tester des effets de marge, où l'adaptation des espèces est très différente (Alberto et al., 2013), du fait de phénomènes éco-évolutifs bien particuliers. Ainsi, cibler des dispositifs où la niche climatique des jardins communs disparaît pour laisser place à un climat non exploré permettrait d'en apprendre davantage sur ces effets de marge. Des dispositifs intégrant de nombreux génotypes différents, très utilisés pour le reboisement, ou alternatifs issus de climats qui

correspondent à la niche climatique future des forêts françaises. D'une manière générale, le déclin observé dans l'installation de tests en jardins communs forestiers depuis 10 ans semble en décalage avec l'énorme potentiel de production de connaissances d'un tel outil, mais aussi l'urgence que constitue l'atténuation et l'adaptation des forêts face au changement climatique.

Les résultats de cette étude sont également à mettre en perspective avec des données pédologiques et écologiques. En effet, les données climatiques seules ne suffisent pas à tirer des conclusions définitives sur la potentielle survie et croissance des populations plantées. Il faut également rappeler que les dispositifs en jardins communs ne testent pas du tout les interactions biotiques qui existent en milieu naturel. Ainsi, de telles recommandations sont à nuancer avec d'autres données pédologiques et écosystémiques, dans l'optique d'une diversification des peuplements.

Conclusion

Ce stage avait pour objectif une analyse descriptive mais aussi prospective de la représentativité climatique du réseau de dispositifs en jardins communs GEN4X. Ce réseau constitue une source d'informations importante au regard des changements climatiques à venir. En effet, les dispositifs expérimentaux qui le constituent représentent un outil efficace pour étudier l'adaptation des espèces, voire des génotypes à un environnement changeant. A travers des mesures et des expérimentations, l'adaptation locale, mais aussi la plasticité phénotypique de ces génotypes peut être qualifiée, et quantifiée. Ces deux processus sont fondamentaux pour l'adaptation des peuplements aux changements climatiques et constituent un enjeu de taille pour la diversification des forêts françaises. Ainsi, connaître davantage sur ces processus d'adaptation permet d'envisager les perturbations qui opéreront sous climats changeants, et ainsi d'adapter la gestion des peuplements forestiers.

Le réseau est lui aussi soumis à un environnement en mutation. En effet, il occupe actuellement une niche climatique bien définie, qui est amenée à changer avec le dérèglement climatique. Ainsi, la représentativité climatique du réseau sur le territoire français va être modifiée, à mesure que les climats évoluent. Cela représente un enjeu majeur de gestion du réseau car certains dispositifs vont ainsi se retrouver dans le futur dans des zones dont le climat sera totalement nouveau en France. A l'inverse, d'autres dispositifs se situeront toujours dans des aires de compatibilité climatique. A travers une analyse prospective, cette étude montre que la représentativité climatique du réseau GEN4X sur le territoire français va être largement modifiée, laissant ainsi une grande partie des dispositifs en climat inconnu. A partir de là, il est important d'ajuster dès maintenant la gestion de ces dispositifs en fonction des connaissances que l'on souhaite produire.

Je me suis attaché dans ce mémoire à démontrer qu'il était possible, avec des données solides, de modéliser de manière robuste l'aire de compatibilité climatique du réseau, et ce, à différentes périodes, avec des outils de type Species Distribution Models. Une approche régionale des dispositifs montre par ailleurs une meilleure efficacité qu'une approche globale du réseau. Il est nécessaire d'améliorer la méthode si l'on souhaite obtenir des résultats fins, mais cela est tout à fait possible. Par ailleurs, une attention particulière a été apportée aux dispositifs dits « méditerranéens » (localisation ou espèces), qui montrent un potentiel de recherche et d'expérimentation important à mesure que ce climat s'étend sur le territoire français. Une approche similaire en prenant en compte cette fois-ci les différents génotypes présents dans les dispositifs GEN4X permettrait d'affiner davantage les résultats. Dans une optique de diversification génétique des forêts françaises, les essences méditerranéennes représentent un atout majeur et leur capacité d'adaptation doit plus que jamais être étudiée.

Bibliographie

Articles

- Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1(1), 95–111. <https://doi.org/10.1111/j.1752-4571.2007.00013.x>
- Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R., & Savolainen, O. (2013). Potential for evolutionary responses to climate change - evidence from tree populations. *Global Change Biology*, 19(6), 1645–1661. <https://doi.org/10.1111/gcb.12181>
- ALLOUCHE, O., TSOAR, A., & KADMON, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223–1232. <https://doi.org/10.1111/J.1365-2664.2006.01214.X>
- Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33(10), 1677–1688. <https://doi.org/10.1111/J.1365-2699.2006.01584.X>
- Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22(1), 42–47. <https://doi.org/10.1016/J.TREE.2006.09.010>
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, 3(2), 327–338. <https://doi.org/10.1111/J.2041-210X.2011.00172.X>
- Barredo, J. I., Mauri, A., Caudullo, G., & Dosio, A. (2018). Assessing Shifts of Mediterranean and Arid Climates Under RCP4.5 and RCP8.5 Climate Projections in Europe. *Pure and Applied Geophysics 2018 175:11*, 175(11), 3955–3971. <https://doi.org/10.1007/S00024-018-1853-6>
- Benito-Garzón, M., & Fernández-Manjarrés, J. F. (2015). Testing scenarios for assisted migration of forest trees in Europe. *New Forests 2015 46:5*, 46(5), 979–994. <https://doi.org/10.1007/S11056-015-9481-9>
- Castellanos-Acuña, D., Lindig-Cisneros, R., & Sáenz-Romero, C. (2015). Altitudinal assisted migration of Mexican pines as an adaptation to climate change <http://www.esajournals.org/doi/pdf/10.1890/ES14-00375.1>. *Ecosphere*, 6(1). <https://doi.org/10.1890/ES14-00375.1>
- Cordell, S., Goldstein, G., Mueller-Dombois, D., Webb, D., & Vitousek, P. M. (1998). Physiological and morphological variation in *Metrosideros polymorpha*, a dominant Hawaiian tree species, along an altitudinal gradient: The role of phenotypic plasticity. *Oecologia*, 113(2), 188–196. <https://doi.org/10.1007/s004420050367>
- De Villemereuil. (2016). *Méthodes pour l'étude de l'adaptation locale et application au contexte de l'adaptation aux conditions d'altitude chez la plante alpine Arabis alpina*. <https://tel.archives-ouvertes.fr/tel-01322336>

- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802–813. <https://doi.org/10.1111/J.1365-2656.2008.01390.X>
- Elith, Jane, & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. <Http://Dx.Doi.Org/10.1146/Annurev.Ecolsys.110308.120159>, 40, 677–697. <https://doi.org/10.1146/ANNUREV.ECOLSYS.110308.120159>
- Elith, Jane, Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1), 43–57. <https://doi.org/10.1111/J.1472-4642.2010.00725.X>
- Guisan, A., Edwards, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, 157(2–3), 89–100. [https://doi.org/10.1016/S0304-3800\(02\)00204-1](https://doi.org/10.1016/S0304-3800(02)00204-1)
- Hao, T., Elith, J., Guillera-Aroita, G., & Lahoz-Monfort, J. J. (2019). A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and Distributions*, 25(5), 839–852. <https://doi.org/10.1111/DDI.12892>
- Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* 2020 7:1, 7(1), 1–18. <https://doi.org/10.1038/s41597-020-0453-3>
- Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12), 1225–1241. <https://doi.org/10.1111/j.1461-0248.2004.00684.x>
- Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12(4), 334–350. <https://doi.org/10.1111/J.1461-0248.2008.01277.X>
- Klausmeyer, K. R., & Shaw, M. R. (2009). Climate Change, Habitat Loss, Protected Areas and the Climate Adaptation Potential of Species in Mediterranean Ecosystems Worldwide. *PLOS ONE*, 4(7), e6392. <https://doi.org/10.1371/JOURNAL.PONE.0006392>
- Langlet, O. (1971). *TAXON* 20(5/6): 653-722 NOVEMBER 1971 TWO HUNDRED YEARS *GENECOLOGY* Summary: Milestones in genecology. <https://about.jstor.org/terms>
- Lefèvre, Fady B, Jean F, Davi H, Pichot C, & Oddou-Muratorio S. (2015). Les processus biologiques de réponse des arbres et forêts au changement climatique : adaptation et plasticité phénotypique. In *Innovations Agronomiques* (Vol. 47).
- Lowe, J., Bernie, D., Bett, P., Bricheno, L., Brown, S., Calvert, D., Clark, R., Karen, Eagle, Edwards, T., Fosser, G., Maisey, P., Mcinnes, R., McSweeney, C., Yamazaki, K., & Belcher, S. (2019). *UKCP 18 Science Overview Report November 2018 (Updated March 2019)*.
- Marchi, M., Castellanos-Acuña, D., Hamann, A., Wang, T., Ray, D., & Menzel, A. (2020). ClimateEU, scale-free climate normals, historical time series, and future projections for Europe. *Scientific Data* 2020 7:1, 7(1), 1–9. <https://doi.org/10.1038/s41597-020-00763-0>
- Martínez-Sancho, E., Rellstab, C., Guillaume, F., Bigler, C., Fonti, P., Wohlgemuth, T., & Vitasse, Y. (2021). Post-glacial re-colonization and natural selection have shaped growth responses of silver fir across Europe. *Science of The Total Environment*, 779, 146393.

<https://doi.org/10.1016/j.scitotenv.2021.146393>

Matyas, C. (1996). Climatic adaptation of trees: rediscovering provenance tests. In *Euphytica* (Vol. 92).

Oleksyn, J., Modrzyński, J., Tjoelker, M. G., Zytkowski, R., Reich, P. B., & Karolewski, P. (1998). Growth and physiology of *Picea abies* populations from elevational transects: Common garden evidence for altitudinal ecotypes and cold adaptation. *Functional Ecology*, *12*(4), 573–590. <https://doi.org/10.1046/j.1365-2435.1998.00236.x>

Pecchi, M., Marchi, M., Burton, V., Giannetti, F., Moriondo, M., Bernetti, I., Bindi, M., & Chirici, G. (2019). Species distribution modelling to support forest management. A literature review. *Ecological Modelling*, *411*, 108817. <https://doi.org/10.1016/J.ECOLMODEL.2019.108817>

Ray, D., Marchi, M., Rattey, A., & Broome, A. (2021). A multi-data ensemble approach for predicting woodland type distribution: Oak woodland in Britain. *Ecology and Evolution*, *11*(14), 9423–9434. <https://doi.org/10.1002/ECE3.7752>

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 595–619. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095646>

Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D., Petr, M., Honkaniemi, J., Lexer, M. J., Trotsiuk, V., Mairota, P., Svoboda, M., Fabrika, M., Nagel, T. A., & O Reyer, C. P. (2017). *Forest disturbances under climate change Europe PMC Funders Group*. 7, 395–402. <https://doi.org/10.1038/nclimate3303>

Thom, D., & Seidl, R. (2016). Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests. *Biological Reviews of the Cambridge Philosophical Society*, *91*(3), 760–781. <https://doi.org/10.1111/brv.12193>

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369–373. <https://doi.org/10.1111/J.1600-0587.2008.05742.X>

Vilà-Cabrera, A., Coll, L., Martínez-Vilalta, J., & Retana, J. (2018). Forest management for adaptation to climate change in the Mediterranean basin: A synthesis of evidence. *Forest Ecology and Management*, *407*, 16–22. <https://doi.org/10.1016/J.FORECO.2017.10.021>

Vitasse, Y., Delzon, S., Bresson, C. C., Michalet, R., & Kremer, A. (2009). *Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden*. <https://doi.org/10.1139/X09-054>

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *PLOS ONE*, *11*(6), e0156720. <https://doi.org/10.1371/JOURNAL.PONE.0156720>

Ouvrage

Guisan, A., Thuiller, W., & Zimmermann, N. (2017). *Habitat Suitability and Distribution Models: With Applications in R* (Ecology, Biodiversity and Conservation). Cambridge: Cambridge University Press. doi:10.1017/9781139028271

Sites web

Centre régional de la propriété forestière Provence-Alpes-Côte d'Azur, s.d., MEDForFUTUR - Adaptation des forêts aux changements climatiques. In : *Site du Centre régional de la propriété forestière Provence-Alpes-Côte d'Azur*. Consulté le 04 septembre 2021. Disponible à l'adresse : <https://paca.cnpf.fr/n/medforfutur-adaptation-des-forets-aux-changements-climatiques/n:3000>

ClimateDT, 2021. In : *Site web du portail ClimateDT*. Consulté le 12 juin 2021. Disponible à l'adresse : <https://ibbr.cnr.it/climate-dt/>

Fady B, IN-SYLVA-France, 2018, Réseau GEN4X (réseau de GENétique FOrestière pour la Recherche et l'eXpérimentation). In : *Site web de IN-SYLVA-France*. Consulté le 06 septembre 2021. Disponible à l'adresse : <https://www6.inrae.fr/in-sylva-france/Presentation>

IGN, 2019, Fiches descriptives des grandes régions écologiques (GRECO) et des sylvoécotons (SER). In : *Site web de l'Inventaire National de l'Information Géographique et Forestière*. Consulté le 25 août 2021. Disponible à l'adresse : <https://inventaire-forestier.ign.fr/spip.php?article773>

IGN, 2019, La diversité des peuplements forestiers In : *Site web de l'Inventaire National de l'Information Géographique et Forestière*. Consulté le 14 septembre 2021. Disponible à l'adresse : <https://inventaire-forestier.ign.fr/spip.php?rubrique78>

WWF, s.d., Terrestrial ecoregions. In : *Site web de WWF*. Consulté le 26 avril 2021. Disponible à l'adresse : <https://www.worldwildlife.org/biome-categories/terrestrial-ecoregions>

Data.gouv.fr, 2010, France : Sylvoécotons. In : Site web de geo.data.gouv. Consulté le 12 juin 2021. Disponible à l'adresse : <https://geo.data.gouv.fr/fr/datasets/Oddaa954951799758fc857dfd9c1d32fb285f536>

Documents techniques, rapports

Bouillie J., ONF, 2021. *Adapter les forêts publiques au changement climatique en Provence-Alpes-Côte d'Azur*. Guide de gestion des peuplements vulnérables et déperissants Sud-PACA. Consulté le 04 septembre 2021. Disponible à l'adresse : file:///C:/Users/grihm/AppData/Local/Temp/ONF_guide%20peuplements%20PACA_2021.pdf

IPCC, 2014 : *Climate Change 2014 : Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Consulté le 25 août 2021. Disponible à l'adresse :
https://www.ipcc.ch/site/assets/uploads/2018/05/SYR_AR5_FINAL_full_wcover.pdf

OST, 2021, *La position scientifique de la France dans le monde et en Europe, 2005-2018*, Hcéres, Paris. Consulté le 29 avril 2021. Disponible à l'adresse :
https://www.hceres.fr/sites/default/files/media/downloads/hceres_ost_positionnement_scientifique_france_edition_2021_1.pdf

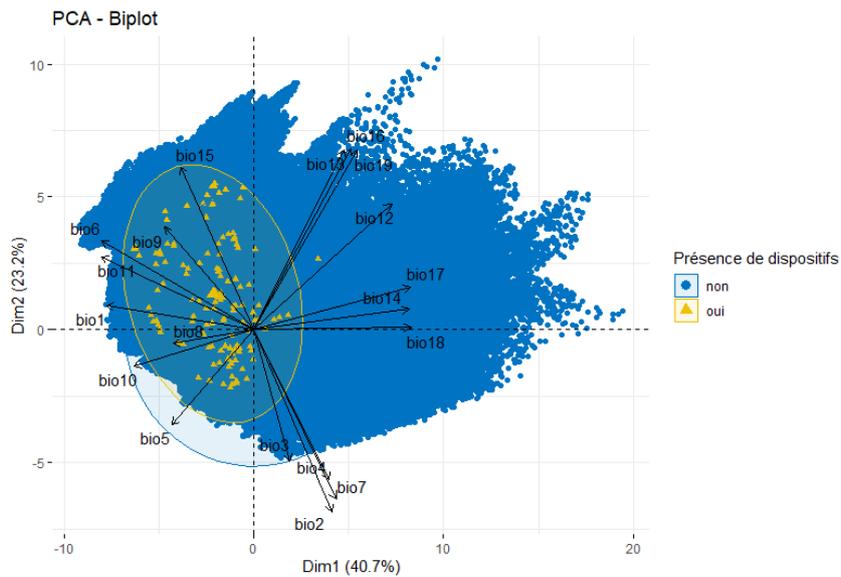
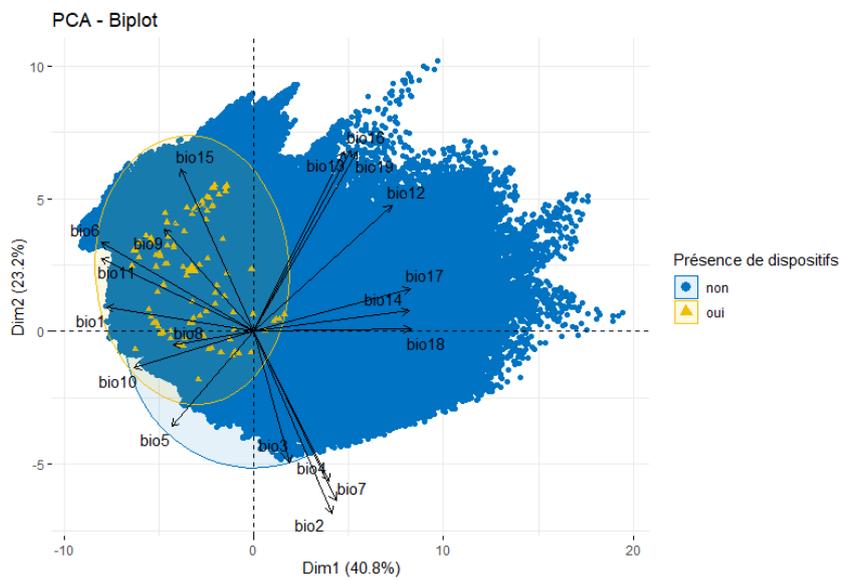
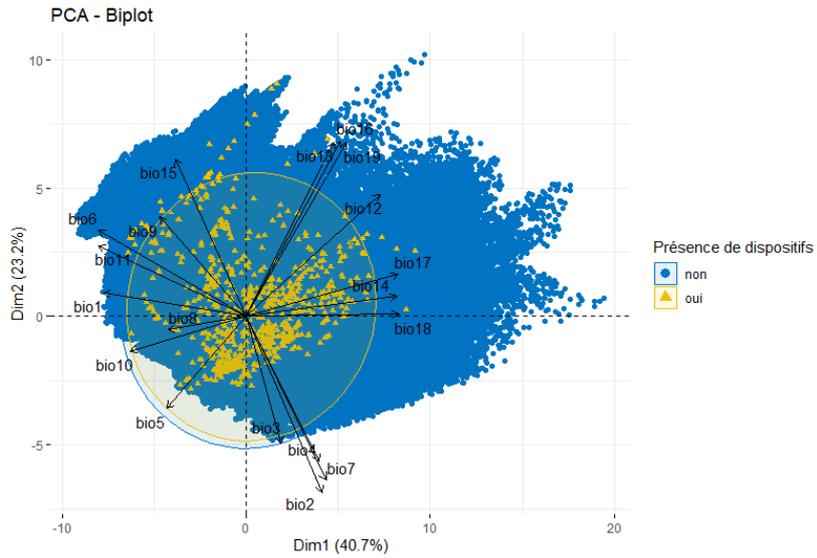
Table des annexes

Annexe I : Graphes des variables et des individus en fonction des 3 jeux de données.

Annexe II : Variables climatiques b1 à b19 de ClimateDT

Annexe III : Liste des types de tests de GEN4X

Annexe IV : Liste des espèces testées dans le réseau GEN4X



Annexe I : Graphes des variables et des individus en fonction des 3 jeux de données. Les individus correspondant aux jeux de données sont en jaune. Les individus correspondant au reste du territoire français sont en bleu. De haut en bas, on retrouve les jeux de données FRANCE, GRECO et ESP.

bio1	42	Mean annual Temperature	Celsius degrees (°C)
bio2	43	Mean Diurnal Range (Mean of monthly (max temp - min temp))	Celsius degrees (°C)
bio3	44	Isothermality (bio2/bio7) (x100)	index
bio4	45	Temperature Seasonality (standard deviation x100)	Celsius degrees (°C)
bio5	46	Max Temperature of Warmest Month	Celsius degrees (°C)
bio6	47	Min Temperature of Coldest Month	Celsius degrees (°C)
bio7	48	Temperature Annual Range (bio5-bio6)	Celsius degrees (°C)
bio8	49	Mean Temperature of Wettest Quarter	Celsius degrees (°C)
bio9	50	Mean Temperature of Driest Quarter	Celsius degrees (°C)
bio10	51	Mean Temperature of Warmest Quarter	Celsius degrees (°C)
bio11	52	Mean Temperature of Coldest Quarter	Celsius degrees (°C)
bio12	53	Mean annual Precipitation	Millimeters (mm)
bio13	54	Precipitation of Wettest Month	Millimeters (mm)
bio14	55	Precipitation of Driest Month	Millimeters (mm)
bio15	56	Precipitation Seasonality (Coefficient of Variation)	index
bio16	57	Precipitation of Wettest Quarter	Millimeters (mm)
bio17	58	Precipitation of Driest Quarter	Millimeters (mm)
bio18	59	Precipitation of Warmest Quarter	Millimeters (mm)
bio19	60	Precipitation of Coldest Quarter	Millimeters (mm)

Annexe II : Variables climatiques bio1 à bio19 de ClimateDT (source : site web de ClimateDT)

Type de dispositif	Effectif
Test de comparaison de descendances	474
Test de comparaison de provenances	416
Test de comparaison de clones	274
Divers	51
Test de comparaison d'espèces	24
Parc à clones	16
Arboretum	13
NA	12
Verger à graines	8
Test Biomasse	2
Test de démonstration	2
Test Hybride	2
Test juvénile	2

Annexe III : Liste des types de tests de GEN4X

Espèce	Effectif	Espèce	Effectif
Abies alba	27	Liriodendron sp.	5
Abies bornmuelleriana	8	Liriodendron tulipifera	43
Abies cephalonica	16	MELANGE ESPECES	9
Abies cilicica	7	MELANGE FEULLUS	8
Abies concolor	6	MELANGE RESINEUX	3
Abies equitro-jani	8	Nothofagus sp.	2
Abies grandis	11	Picea abies	100
Abies lowiana	5	Picea mariana	1
Abies Nobilis	1	Picea sitchensis	15
Abies nordmaniana	18	Picea sitka	1
Abies numidica	3	Picea sp.	3
Abies pinsapo	4	Pinus brutia	10
Abies procera	3	Pinus contorta	6
Abies sp.	9	Pinus eldarica	10
Acacia melanoxylon	4	Pinus halepensis	10
Acer pseudoplatanus	3	Pinus nigra	58
Alnus cordata	7	Pinus pinaster	57
Alnus rubra	1	Pinus pinea	8
Alnus glutinosa	4	Pinus ponderosa	4
Alnus sp.	2	Pinus radiata	4
Betula sp.	1	Pinus sp.	2
Calocedrus sp.	1	Pinus sylvestris	40
Carpinus sp.	1	Pinus taeda	1
Cedrus atlantica	35	Platanus sp.	2
Cedrus brevifolia	12	Populus alba	15
Cedrus deodara	5	Populus deltoides	27
Cedrus libani	32	Populus leuce	69
Cedrus sp.	10	Populus maximowiczii	3
Cryptomeria japonica	7	Populus nigra	6
Cupressocyparis sp.	1	Populus sp.	10
Cupressus arizonica	12	Populus tremula var. glandulosa	15
Cupressus atlantica	2	Populus trichocarpa	24
Cupressus dupreziana	7	Prunus avium	88
Cupressus sempervirens	30	Prunus avium L.	1
Cupressus sp.	5	Prunus serotina	6
Eucalyptus sp.	18	Pseudotsuga menziesii	127
Fagus sylvatica	26	Quercus petraea	19
Fraxinus excelsior	17	Quercus robur	2
Juglans hybride	2	Quercus rubra	34
Juglans nigra	2	Quercus sp.	19
Juglans regia	12	Quercus suber	1
Juglans sp.	1	Robinia sp.	1
Larix decidua	142	Salix sp.	2
Larix kaempferi	119	Sorbus domestica	17
Larix laricina	2	Thuja plicata	3
Larix sp.	4	Thuja sp.	11



RIHM Gaspard, 2021, Représentativité climatique du réseau de dispositifs en jardins communs GEN4X et particulièrement des dispositifs méditerranéens sur le territoire français dans un contexte de réchauffement global, 39 pages, mémoire de fin d'études, VetAgro Sup, 2021.

STRUCTURE D'ACCUEIL ET INSTITUTIONS ASSOCIEES:

- ♦ INRAE URFM, UEFM

ENCADRANTS :

- ♦ Maître de stage : FADY, Bruno (Directeur de recherches, URFM, INRAE)
- ♦ Tuteur pédagogique : TARDIF, Antoine

OPTION : Agriculture Environnement Santé Territoire

RESUMÉ

Les forêts françaises font face aujourd'hui au réchauffement climatique et leur adaptation est un enjeu majeur de la gestion de cette ressource. La recherche en sciences forestières est plus que jamais primordiale. Le réseau GEN4X, qui regroupe plus d'un millier de dispositifs en jardins communs répartis sur le territoire Français métropolitain, fait partie des infrastructures de recherche qui permettent l'étude de l'adaptation des essences forestières aux changements climatiques. Ce réseau, lui aussi, doit faire face aux changements environnementaux.

Ce mémoire présente le travail d'analyse descriptive et prospective du réseau GEN4X, qui cherche à mettre en évidence la représentativité climatique du réseau sur le territoire français. Pour cela, des outils de modélisation type Species Distribution Models et notamment la plateforme BIOMOD ont permis de mettre en évidence des tendances concernant la niche climatique du réseau entier. Une attention particulière a été portée aux dispositifs situés en région méditerranéenne, et comportant des essences méditerranéennes. En effet, ces dispositifs permettent de décrire plus finement l'évolution de la représentativité climatique du réseau et possèdent un important potentiel de diversification génétique des forêts françaises.

Mots clés : Jardins communs – niche climatique – Species Distribution Models – BIOMOD – Changement climatique – Adaptation - Forêt

ABSTRACT

French forests are currently facing global warming and their adaptation is a major issue in the management of this resource. Research in forest sciences is more important than ever. The GEN4X network, which includes more than a thousand common garden experiments spread over the French metropolitan territory, is part of the research infrastructure that allows the study of the adaptation of forest species to climate change. This network must also face environmental changes.

This master thesis presents the descriptive and prospective analysis of the GEN4X network, which seeks to highlight the climatic representativeness of the network over the French territory. For this purpose, modeling tools such as Species Distribution Models and particularly BIOMOD platform have been used to highlight trends concerning the climatic niche of the entire network. Particular attention was paid to the systems located in the Mediterranean region and including Mediterranean species. Indeed, these systems allow a more detailed description of the evolution of the climatic representativeness of the network and have an important potential for the genetic diversification of French forests.

Mots clés : Common garden experiments – climatic niche – Species Distribution Models – BIOMOD – Climate Change – Adaptation – Forest

