DMP du projet "IMpacts des PRocessus mIcroclimatiques sur la redistributioN de la biodiversiTé forestière en contexte de réchauffement du macroclimat"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - Modèle de PGD (français)" fourni par Agence nationale de la recherche (ANR).

Renseignements sur le plan

Titre du plan	DMP du projet "IMpacts des PRocessus mIcroclimatiques sur la redistributioN de la biodiversiTé forestière en contexte de réchauffement du macroclimat"
Livrable	
Version	Version finale
Objet/périmètre du plan	L'objet du présent PGD est de présenter la manière dont ont été gérées les données acquises au cours du projet IMPRINT et de présenter une partie des données déjà partagées via l'entrepôt de données InDoRES.
Domaines de recherche (selon classification de l'OCDE)	1.5 Earth and related environmental sciences, 4.1 Agriculture, forestry, and fisheries, 1.6 Biological sciences
Langue	fra
Date de création	2020-01-29
Date de dernière modification	2024-11-06
Identifiant	PGD-IMPRINT
Type d'identifiant	DOI
Licence	**Nom** Creative Commons Attribution 4.0 International **URL** http://spdx.org/licenses/CC-BY-4.0.json
Documents (publications, rapports, brevets, plan expérimental....), sites web associés	• Lenoir, Gril, Durrieu, Horen, Laslier, Lembrechts, Zellweger, Alleaume, Brasseur, Buridant, Dayal, De Frenne, Gallet-Moron, Marrec, Meeussen, Rocchini, Van Meerbeek & Decocq (2022) Unveil the unseen: Using LiDAR to capture time-lag dynamics in the herbaceous layer of European temperate forests. Journal of Ecology, vol. 110, issue 2, pages 282-300 : 10.1111/1365-2745.13837 • Gril, Spicher, Greiser, Ashcroft, Pincebourde, Durrieu, Nicolas, Richard, Decocq, Marrec & Lenoir (2023) Slope and equilibrium: A parsimonious and flexible approach to model microclimate. Methods in Ecology and Evolution, vol. 14, issue 3, pages 885-897 : 10.1111/2041-210X.14048 • Gril, Laslier, Gallet-Moron, Durrieu, Spicher, Le Roux, Brasseur, Haesen, Van Meerbeek, Decocq, Marrec & Lenoir (2023) Using airborne LiDAR to map forest microclimate temperature buffering or amplification. Remote Sensing of Environment, vol. 298, n°113820 : 10.1016/j.rse.2023.113820

- Laslier, Lalechère, Gril, Ferrer & Lenoir (2023) Mapping forest microclimates using Sentinel 1 radar data: insights and complementarity with LiDAR and optical data. Remote Sensing for Agriculture, Ecosystems, and Hydrology XXV, vol. 12727, 127270C : 10.1117/12.2679400
- Gril, Spicher, Vanderpoorten, Vital, Brasseur, Gallet-Moron, Le Roux, Decocq, Lenoir & Marrec (2024) Ecological indicator values of understorey plants perform poorly to infer forest microclimate temperature. Journal of Vegetation Science, vol 35, n°e13241 : 10.1111/jvs.13241
- Brusse, Lenoir, Boisset, Spicher, Dubois, Caro & Marrec (2024) Mechanistically mapping near-surface temperature in the understory of temperate forests: A validation of the microclima R package against empirical observations. Agricultural and Forest Meteorology, vol. 346, n°109894 : 10.1016/j.agrformet.2024.109894
- Michalet, Carcaillet, Delerue, Domec & Lenoir (2024) Assisted migration in a warmer and drier climate: less climate buffering capacity, less facilitation and more fires at temperate latitudes? Oikos, vol 8, n°e10248 : 10.1111/oik.10248
- Gril, Spicher, Vanderpoorten, Gallet-Moron, Brasseur, Le Roux, Laslier, Decocq, Marres, Lenoir (2024) The affinity of vascular plants and bryophytes to forest microclimate buffering. Journal of Ecology, Early View : 10.1111/1365-2745.14425
- Site web du projet IMPRINT : https://microclimat.cnrs.fr

Renseignements sur le projet

Titre du projet

IMpacts des PRocessus mIcroclimatiques sur la redistributioN de la biodiversiTé forestière en contexte de réchauffement du macroclimat

Acronyme

IMPRINT

Résumé

Les modèles de niche, couramment utilisés pour prédire la redistribution du vivant en contexte de réchauffement climatique, sont habituellement calibrés sur les températures synoptiques (i.e. macroclimat) mais ignorent les températures ressenties (i.e. microclimat). Pourtant, le microclimat ressenti par de nombreux organismes peut être très différent du macroclimat régional, surtout au sein des écosystèmes forestiers dont la sylviculture peut servir de médiateur entre microclimat et macroclimat. Des progrès récents permettent aujourd'hui de spatialiser le microclimat à très fine résolution en combinant des mesures microclimatiques locales avec des variables environnementales issues de la télédétection 3D à haute définition (cf. technologie LiDAR). Néanmoins, cette spatialisation du microclimat reflète rarement la dynamique des fluctuations interannuelles à long terme du climat (échelle de la décennie) mais plutôt les variations météorologiques intra-annuelles (météo plutôt que climat), témoignant d'un manque de données sur le suivi microclimatique à long terme. Afin de lever ce verrou scientifique, nous proposons d'utiliser de manière combinée : (i) des micro-capteurs de température et d'humidité afin de mesurer en continue le microclimat de surface ; (ii) des images 3D à très haute définition issues de la technologie LiDAR aéroportée ; et (iii) des données de suivi climatique à long terme issues d'un réseau national de postes météo

permanents. Le projet IMPRINT s’appuiera sur cette combinaison d’informations pour non seulement quantifier et modéliser les processus sous-jacent au microclimat sur de grandes étendues spatiales et à une résolution spatiotemporelle fine, mais aussi et surtout pour reconstruire les tendances d’évolution du microclimat sous-couvert forestier et donc l’intensité du réchauffement ressentie au sein de ces écosystèmes. L’objectif ultime du projet IMPRINT est d’utiliser cette reconstruction à long terme du microclimat sous-couvert forestier dans des modèles de niche afin de montrer comment le microclimat affecte les prédictions de redistribution de la biodiversité forestière en contexte de réchauffement et ce par rapport aux modèles de niche actuels basés sur le macroclimat. Nous supposons que la prise en compte du microclimat sous-couvert forestier dans les modèles de niche va non seulement impliquer un décalage de la niche climatique des espèces vers des conditions plus fraiches par rapport à un modèle de niche basé sur le macroclimat mais également que cela favorisera un retard de réponse en contexte de réchauffement à cause du découplage entre température sous-couvert et hors-couvert (un réchauffement de 1°C hors-couvert implique un réchauffement moindre sous-couvert sur le même laps de temps). Pour tester ces hypothèses suggérant que les processus microclimatiques peuvent agir comme un frein dans la dynamique de redistribution du vivant, nous nous focaliserons sur les plantes vasculaires et les arthropodes des forêts tempérées feuillues, deux groupes taxonomiques jouant un rôle important sur le fonctionnement des cycles biogéochimiques et sur la provision de certains services écosystémiques. La principale nouveauté de ce projet repose sur le fait que nous n’allons pas modéliser le microclimat directement comme c’est le cas actuellement mais plutôt indirectement en modélisant d’abord les processus qui relient le microclimat au macroclimat et en utilisant ensuite une fonction de transfert basée sur ces processus pour pouvoir reconstruire le microclimat passé à partir d’observations issues de postes météo permanents. Les résultats et retombés de ce projet ont la capacité de remettre en cause les prédictions actuelles des modèles de niche et de proposer des stratégies de gestion forestière permettant d’agir sur le microclimat sous-couvert forestier et ainsi de limiter les effets néfastes du réchauffement climatique sur la biodiversité des forêts tempérées feuillues.

Sources de financement

- Agence nationale de la recherche (ANR) : ANR-19-CE32-0005-01
- Agence Nationale de la Recherche : ANR-19-CE32-0005
- Agence Nationale de la Recherche : ANR-19-CE32-0005
- Agence Nationale de la Recherche : ANR-19-CE32-0005

Date de début 2019

Date de fin 2024-03-31

Produits de recherche :

1. Positions géographiques et informations générales sur les placettes de suivi installées en forêts domaniales de Mormal, Blois et l'Aigoual (Jeu de données)
2. Séries temporelles de températures microclimatiques sur les placettes IMPRINT installées en forêts domaniales de Mormal, Blois et l'Aigoual (Jeu de données)
3. Données lidar du survol des trois forêts domaniales du projet IMPRINT (Mormal, Blois, Aigoual) (Jeu de données)

4. Inventaires dendrométriques des placettes IMPRINT installées en forêts domaniales de Mormal, Blois et l'Aigoual (Jeu de données)
5. Inventaires floristiques des placettes IMPRINT installées en forêts domaniales de Mormal, Blois et l'Aigoual (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
Lenoir Jonathan - https://orcid.org/0000-0003-0638-9582	Ecologie et dynamique des systèmes anthropisés - 201220280W	• Coordinateur du projet • Personne contact pour les données (Lidar IMPRINT, Placettes IMPRINT, Microclimats IMPRINT, Dendrométrie IMPRINT, Flore IMPRINT) • Responsable du plan

DMP du projet "IMpacts des PRocessus mIcroclimatiques sur la redistributioN de la biodiversiTé forestière en contexte de réchauffement du macroclimat"

---

1. Description des données et collecte ou réutilisation de données existantes

1a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

Les données recueillies et produites sur la durée du projet IMPRINT et décrites dans ce plan de gestion des données sont issues de trois forêts domaniales (FD) en France (FD de l'Aigoual, FD de Blois et FD de Mormal) et concernent quatre types de données : (i) des **données microclimatiques recueillies in-situ** (p.ex. enregistrements horaires de températures de l'air ou du sol) ; (ii) des **données d'observations du milieu réalisées in-situ** (p.ex. données de localisation géographique, relevés dendrométriques et inventaires botaniques) ; (iii) des **données issues de prélèvements réalisés in-situ puis analysés en laboratoire** (p.ex. prélèvements d'échantillons de sols pour mesure de pH du sol en laboratoire ou pose de pièges Barber pour collecter les arthropodes du sol à identifier à postériori au laboratoire) ; et (iv) des **données issues de la télédétection** (p.ex. nuages de points 3D issus de la technologie LiDAR). Pour chacune des trois FD étudiées, un réseau de 60 placettes de suivies a été installé et suivi sur les quatre années du projet (2020-2023). Voir le site web du projet IMPRINT pour plus d'information sur la répartition géographique des données à l'échelle nationale : https://microclimat.cnrs.fr/zones-detude/.

(i) Les données microclimatiques (p.ex. températures de l'air et du sol et humidité du sol) ont été recueillies à l'aide de capteurs de données environnementales posés directement in-situ, au sein des trois FD étudiées, par les scientifiques impliqués dans le projet. Deux types de capteurs ont été utilisés et achetés sur le budget du projet : (i) des capteurs HOBO UA-001-64 posé à 1 m au dessus du sol pour mesurer la température de l'air au pas horaire et (ii) des capteurs HOBO UA-001-08 posés à 8 cm sous la surface du sol pour mesurer la température du sol au pas de 2 heures. Les capteurs ont été installés dès la première année du projet (été 2020). Une autre série de capteurs de la marque TOMST (TMS4) ont été utilisés et achetés sur le budget du projet pour mesurer la température à 10 cm au dessus du sol, au niveau du sol et à 8 cm sous la surface du sol ainsi que l'humidité relative du sol à 8 cm sous la surface du sol avec une fréquence d'un enregistrement toute les 15 mins pour les quatre paramètres microclimatiques sus-mentionnées. Ces capteurs ont été envoyés dès la première année du projet (2020) par colis postal, accompagnés d'un protocole d'installation in-situ, aux agents forestiers du Réseau National de suivi à long terme des ECOsystèmes FORestiers (RENECOFOR) de l'Office National des Forêts (ONF). En effet, une partie du projet IMPRINT s'est appuyé sur le dispositif expérimental créé et géré par le réseau RENECOFOR pour la collecte et la mise à disposition de données existantes. Les données microclimatiques ont été déchargées au moins une fois par an (2021 et 2022), lors des campagnes de terrain ou au cours des réunions annuelles du réseau RENECOFOR, à l'aide des logiciels dédiés HOBOware et Lolly Manager puis analysées sous le logiciel libre R de traitements statistiques. Il est à noter que suite à la première année de mesure des conditions microclimatiques sous couvert forestier nous avons fait face à un problème majeur de dysfonctionnement d'une grande partie des capteurs installés sur les trois FD étudiées dans le cadre du projet, vraisemblablement lié à l'utilisation de piles défectueuses ou de mauvaises qualités. Une partie des données microclimatiques, notamment les données de températures de l'air et du sol mesurées sur les trois FD étudiées dans le cadre du projet, a déjà fait l'objet d'un dépôt sur l'entrepôt de données InDoRES (10.48579/PRO/ZQBQUW). Les données microclimatiques issues des placettes RENECOFOR de l'ONF n'a pas encore fait l'objet d'un dépôt sur InDoRES car en cours de valorisation.

(ii) Les données d'observations du milieu (p.ex. localisation des placettes de mesure in-situ, relevés dendrométriques et botaniques, etc.) ont été collectées par chaque scientifique impliqué dans le projet et ce sur les les trois FD étudiées dans le projet. La *"provenance des données"* ou localisation géographique (latitude, longitude, altitude) des placettes de mesures a été documentée dès l'installation des placettes de suivie du microclimat, la première année (été 2020), à l'aide d'un GPS Trimble et d'une antenne pour une meilleure précision. Les données dendrométriques (p.ex. le diamètre des arbres à 1,30 m) de structure du peuplement forestier ont été mesurées in-situ dès l'installation des placettes en 2020, avec prises de photographies hémisphériques pour quantifier le degré d'ouverture du couvert forestier sur chaque placette. La composition en espèces végétales dans les différentes strates de végétation (muscinale, herbacée, arbustive et arborée) a été inventoriée sur des placettes de surface fixe (400 m2) au cours de la deuxième année du projet (été 2021). La saisie et la numérisation des données d'observation du milieu a été faite directement sur le terrain à l'aide d'un tableur Excel et d'une tablette endurcie et étanche pour faire face aux aléas climatiques liés aux conditions du terrain. L'analyse des données a été réalisée à l'aide du logiciel libre R. Les données de localisation des placettes, les données dendrométriques issues des inventaires forestiers ainsi que les données d'inventaires floristiques (flore vasculaire et bryophytes) des trois FD étudiées sont déjà disponibles sur InDoRES (10.48579/PRO/RSBT3C, 10.48579/PRO/AIXCGM, 10.48579/PRO/BJ9MT6).

(iii) Les données issues de prélèvements réalisés in-situ et analysés en laboratoire comprennent la prise d'échantillons de sol et la pose de pièges Barber pour déterminer la composition des communautés d'arthropodes du sol. Les prélèvements de sol ont été effectués sur les trois FD étudiées et ce dès la première année (été 2020), lors de l'installation des placettes de suivi des conditions microclimatiques. Les échantillons de sol ont ensuite été analysés en laboratoire pour mesurer le pH du sol notamment. Les pièges Barber ont été posés in-situ (et récupérés 15 jours plus tard) par l'équipe de scientifique impliqués dans le projet mais uniquement sur deux des trois FD étudiées dans le cadre du projet : les FD de Bois et de Mormal. Deux campagnes de prélèvements étant nécessaire, au printemps et à l'été, il n'a pas été possible de réaliser ces inventaires de la faune carabique sur les trois FD. D'autre part, la pose et la dépose des pièges Barber a été étalée sur deux années : été 2021 pour la FD de Mormal et été 2022 pour la FD de Blois. Les identifications taxonomiques ont été effectué en laboratoire au cours de la troisième année du projet (2022) par l'équipe

d'entomologues impliqués dans le projet et spécialistes du groupe des carabes. L'identification des individus piégés a été faite, autant que faire se peut, au rang de l'espèce, du genre ou de la famille, suivant les compétences en interne. Nous avons également fait appel, via une prestation de service, à des spécialistes de certains genres taxonomiques ainsi qu'à des spécialistes d'autres groupes taxonomiques comme les araignées que nous avons également piégés dans les pot pièges Barber. Les données de composition taxonomique ont été saisies dans un tableur Excel avant d'être analysées sous le logiel libre R. Ces données n'ont pas encore fait l'objet d'un dépôt sur InDoRES mais les données de pH du sol ont déjà été déposées sur InDoRES (10.48579/PRO/RSBT3C).

(iv) Les données issues de la télédétection (nuages de points 3D LiDAR) ont été collectées par un prestataire externe (Société ALTOA) au projet lors d'un survol des trois FD étudiés. Cette campagne d'acquisition des données LiDAR a eu lieu lors de la seconde campagne de terrain et pendant la période pleine feuille (été 2021) afin de bien caractériser le couvert végétal des trois FD étudiées qui sont dominées par des essences feuillues (chênes et hêtres). Les images LiDAR ont été traitées et analysées sous le logiciel ArcGIS ainsi que sous le logiciel R. Une partie des données LiDAR a déjà été déposée sur InDoRES (10.48579/PRO/XH62TC).

En ce qui concerne les données existantes et qui ont été utilisées au cours du projet, celles-ci comprennent pour l'essentiel des données forestières (parcellaire, type de peuplement, âge du peuplement, gestion, etc.) collectées par l'ONF mais aussi des données climatiques (températures et précipitations) en libre accès (bases de données WorldClim, CHELSA et TerraClimate) ou bien issues du réseau Météo France. Nous avons établis plusieurs conventions d'échanges de données, notamment avec l'ONF et Météo France afin d'établir les éventuelles restrictions à la réutilisation des données préexistantes. Pour les données en libre accès, les restrictions éventuelles d'utilisation sont consultables sur les différents sites Internet dédiés à ces données. Les données cartographiques ont été traitées et analysées sous le logiciel ArcGIS ou sous le logiciel R.

1b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

Les **protocoles** liés à la collecte des données, qu'il s'agissent des données collectées sur le terrain (inventaires dendrométriques, floristiques et faunistiques), des données collectées en laboratoire (identification taxonomique à la loupe binoculaire) ou bien des données collectées par un prestataire de service (survol LiDAR) seront tous conservés soit sous forme papier (dans les cahiers de laboratoire ou dans un cahier des charges auprès du prestataire assurant l'acquisition des images LiDAR) soit au format élétronique (.docx, .pdf).

Les données qui ont été acquises au cours du projet sont pratiquement toutes de type **numérique** (tableurs de données et géodatabases). Quelques images (p.ex. photos hémisphériques ou photos traditionelles) ou vidéos ont également été acquises au cours du projet afin d'illustrer un site web ou un blog dédié au projet. En ce qui concerne les données textuelles, celles-ci concernent uniquement les protocoles d'acquisition des données ou bien les fichiers de métadonnées.

Une grande partie des données issues des trois FD étudiées dans le projet IMPRINT sont déjà disponibles sur l'entrepôt de données InDoRES avec création d'un compte dédié au projet IMPRINT (https://data.indores.fr/dataverse/imprint). Les données déjà déposées sur InDoRES ont déjà été valorisées dans des articles scientifiques et sont organisées en 5 blocs (le bloc placettes, le bloc microclimats, le bloc LiDAR, le bloc dendrométrie et le bloc flore) de fichiers avec pour chacun des 5 blocs, un ou plusieurs fichiers par FD. Le choix d'organiser les données par bloc de données et par FD a été fait pour faciliter l'utilisation des données par les utilisateurs qui ne seraient intéressés que par une partie des données seulement, comme les données issues d'une seule FD en particulier et ainsi éviter de devoir télécharger la totalité des données pour au final n'utiliser qu'une partie de celles-ci. Les données non encore déposées sur InDoRES seront déposés une fois les données valorisées sous la forme de publications scientifiques. C'est le cas notamment des données d'inventaires des arthropodes du sol ou encore des données microclimatiques issues des placettes RENECOFOR qui n'ont pas encore fait l'objet d'une publication scientifique.

Les données de localisation des placettes (10.48579/PRO/RSBT3C) initialement stockées au format tableur électronique (.txt, .csv, .xlsx) ont été transformées au format tabulaire (.tab) et au format shapefile (.shp), soit deux fichiers par FD de moins de 10 Ko par fichier. Un fichier readme de 2.4 Ko est également mis à disposition au format texte (.txt). Les séries temporelles de température de l'air et du sol issues des capteurs HOBO peuvent être enregistrés dans plusieurs formats puis traités par n'importe quel logiciel de traitement de données. Ces données ont initialement été enregistrées au format tableur électronique (.txt, .csv, .xlsx) puis analysées au format R (.RData) dans le logiciel RStudio pour finalement être stockées au format tabulaire (.tab) dans l'entrepôt InDoRES (10.48579/PRO/ZQBQUW). Le volume des données microclimatiques qui ont été générées pendant la durée du projet est de 169.6 Mo pour la FD de l'Aigoual, 159.6 Mo pour la FD de Blois et 154.0 Mo pour la FD de Mormal. Un fichier readme de 1 Ko est également mis à disposition au format texte (.txt). Les données d'observations du milieu comme les données issues des inventaires dendrométriques et floristiques, initialement enregistrées au format tableur électronique (.txt, .csv, .xlsx), ont également été convertis et stockées au format tabulaire (.tab) dans l'entrepôt InDoRES (10.48579/PRO/AIXCGM, 0.48579/PRO/BJ9MT6). Le volume de données floristiques varie entre 80 et 90 Ko par FD. De même, les données d'analyses de sol (pH, C/N, etc.), acquises en laboratoire, ont initialement été enregistrées au format tableur électronique (.txt, .csv, .xlsx) puis convertis au format tabulaire (.tab). Les données LiDAR sont stockées au format dédié (.las), avec un volume de données pouvant dépasser plusieurs To par FD. L'ensemble de ces données brutes sont pour l'instant stockées sur un serveur au sein de l'unité EDYSAN. Seule une partie des nuages de points 3D bruts de chaque FD a été extraite, sur un rayon de 20 m autour du centre de chacune des 60 placettes par FD, et déposé sur InDoRES (10.48579/PRO/XH62TC).

Dans la mesure du possible, des formats standards (.txt, .tab, .las, .shp) et ouverts ont été privilégiés à des fins de partage et de réutilisation. Une géodatabase a été crée sous le logiciel ArcGIS par la géomaticienne impliquée dans le projet (Emilie Gallet-Moron) afin de pouvoir gérer les données du projet IMPRINT et de pouvoir extraire une partie des données suivant les besoins du projet IMPRINT. C'est d'ailleurs Emilie Gallet-Moron qui a extrait les données de la géodatabase qui ont ensuite été téléversées sur l'entrepôt InDoRES.

---

2. Documentation et qualité des données

2a. Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?

Chacun des 5 blocs de données que nous avons téléversé sur l'entrepôt InDoRES est accompagné d'un fichier de métadonnées téléchargeable au format DDI. Ce document liste les auteurs et fourni une description détaillée (en anglais) des données ainsi que des publications connexes associées. Quand cela est nécessaire, certains fichiers sont également accompagnés d'un fichier texte (.txt) de type "*Read_me.txt*" indiquant les métadonnées telles que la liste des noms (header) des variables enregistrées, avec pour chaque variable (header), le type d'appareillage utilisé pour la prise de mesure, les conditions dans lesquelles les données ont été collectées, l'unité de la variable mesurée ou toutes autres informations permettant une réutilisation des données. La langue utilisée est l'anglais pour une plus grande réutilisation potentielle des données collectées. Le standard de métadonnées utilisé dans le cadre du projet est le standard "*Ecological Matadata Language*" (EML) avec une implémentation possible sous le logiciel R grace au package EML.

La convention de nommage des fichiers ou dossiers de données est celle recommandée sur le site DoRANum en utilisant si possible la date, au format "*AAAAMMJJ*", de la dernière version du fichier de données, sans utiliser d'accents ou de caractères spéciaux mais le symbole "_" ou "*underscore*" en tant que séparateur de mots dans le nom du fichier et enfin en affichant l'élément important en premier dans le nom (p.ex. "*20200218_DMP_IMPRINT.docx*"). Dans le cas de documents pouvant être amenés à évoluer, nous préciserons la version dans le nom du fichier (p.ex. "*20200218_DMP_IMPRINT_v01.docx*").

Pour certaines des publications connexes associées aux données du projet IMPRINT, une documentation détaillée des données utilisées ainsi que des analyses effectuées à partir des données est disponible au format RMarkdown (.rmd) généré sous le logiciel libre RStudio. A partir du code brut (.rmd), plusieurs formats de sorties sont possibles (.html, .pdf, .docx) pour documenter les données en terme de : (i) méthodologie utilisée pour acquérir ou analyser la donnée ; (ii) nature des variables utilisées ; (iii) d'unités de mesure utilisées ; ou bien (iv) de distribution de la donnée. Des sorties graphiques permettant une meilleure visualisation des données et de leurs distributions peuvent être intégrées directement dans le document final, avec ou sans le code brute utilisé pour générer ces sorties graphiques.

2b. Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?

De manière générale, les bonnes pratiques de terrain et de laboratoire ont été appliquées pour le contrôle et la qualité des données, que ce soit lors de la collecte de données in-situ (en forêt) ou bien en laboratoire sur des échantillons prélevés in-situ, sur le terrain. Le protocole d'échantillonnage des conditions microclimatiques le long d'un gradient d'ouverture-fermeture de la canopée forestière a permis de couvrir l'ensemble de la variabilité disponible. Comme prévu initialement, trois FD à dominante feuillues ont été échantillonnées (FD de l'Aigoual, FD de Blois et FD de Mormal). Au sein de chaque FD, 60 placettes (un sous-ensemble de placettes de calibration et un sous-ensemble de placettes de validation) a été (i) équipées de capteurs pour enregistrer les conditions microclimatiques sous-couvert forestier et (ii) inventoriées sur le plan de la biodiversité végétale et animale (cf. communautés d'arthropodes du sol). Nous avons également équipé une partie du réseau RENECOFOR (les placettes de forêts feuillues notamment) avec un jeu apparié de sondes TMS4 : (i) une sonde TMS4 sous-couvert forestier au sein de la placette de suivie du réseau RENECOFOR et (ii) une autre sonde TMS4 placée hors-couvert forestier et au sein d'une station météo située à proximité de la forêt, lorsque ce type de station météo est disponible. Ce jeu de données microclimatiques recueillies sur l'ensemble du territoire national est très complémentaire du précédent qui est focalisé sur trois FD équipées d'une plus grande densité de capteurs microclimatiques.

Concernant la qualité et la conformité de la collecte des données microclimatiques, une phase d'intercalibration des capteurs HOBO UA-001-08, HOBO UA-001-64 et TMS4 en conditions contrôlées a été réalisée. L'installation des capteurs de température (air et sol) et d'humidité relative (sol), in-situ, a suivi un protocole standardisé pour l'ensemble des sites étudiés. Concernant les données saisies sur le terrain (inventaires dendrométriques et floristiques) ou en laboratoire (détermination des espèces de carabes et analyses de sol éventuellement), elles ont été validées par l'ensemble des scientifiques impliqués dans le projet.

---

3. Stockage et sauvegarde pendant le processus de recherche

3a. Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?

Sur toute la durée du projet, les données et les métadonnées générées par le projet IMPRINT ont été hébérgées sur les différents postes de travail chiffrés des membres du projet. Pour les données les plus volumineuses (données LiDAR), elles ont été stockées et sauvegardées sur plusieurs disque dur externe, également chiffrés, ainsi que sur le serveur NAS de l'unité de recherche EDYSAN (UMR 7058 CNRS). Pour les données les moins volumineuses (cf. les données microclimatiques et les données issues des inventaires dendrométriques, floristiques et faunistiques), ainsi que toutes les informations relatives au projet, et ce pour faciliter l'échange

d'information entre les membres du projet, nous avons fonctionné avec l'utilisation d'un dossier partagé et intitulé "*IMPRINT*" sur la plateforme de sauvegarde et de partage sécurisé du CNRS : myCoRe. Ce dossier a récemment été migré sur la nouvelle plateforme de sauvegarde et de partage sécurisé du CNRS : sDrive. Ce dossier n'est désormais plus partagé, depuis la fin du projet et sa migration sur la plateforme sDrive.

Des sauvegardes (backup) régulières (tous les 6 mois au moins) de l'ensemble des données et métadonnées du projet (inclues les données LiDAR) ont également été effectuées et enregistrées sur le serveur NAS de l'unité de recherche EDYSAN (UMR 7058 CNRS). A noter également que les données LiDAR ont été partagées avec chaque agence territoriale de l'ONF concernée par le projet (FD de l'Aigoual, FD de Blois et FD de Mormal). Cela assure une sauvegarde supplémentaire de ces données et permet une meilleure valorisation des données acquises au cour du projet. L'agence territoriale de la FD de Mormal a notamment profité du levé LiDAR réalisé en juin 2021 pour effectuer une campagne de relevés dendrométriques sur l'ensemble de la FD de Mormal dans le but de calibrer un modèle d'estimation de la biomasse.

3b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?

En cas d'incident, les données et métadonnées les moins volumineuses pourront être récupérées facilement à partir de la plateforme de sauvegarde et de partage sécurisé du CNRS (sDrive). Pour ce qui est des données LiDAR, plus volumineuses, elles pourront être récupérées à partir de la sauvegarde effectuée sur le NAS de l'unité de recherche EDYSAN (UMR 7058 CNRS). Nous avons également effectué une copie des données LiDAR et de l'ensemble de la géodatabase dur projet IMPRINT sur un cloud (espace NextCloud) avec un espace de stockage conséquent et en accord avec le service de la DISI de l'Université de Picardie Jules Verne (UPJV).

Sur toute la durée du projet, l'ensemble des scientifiques impliqués dans le projet a eu accès aux données et métadonnées les moins volumineuses via la plateforme de sauvegarde et de partage sécurisé du CNRS (myCoRe). Pour les données plus volumineuses, l'accès s'est effectué au cas par cas et suivant les besoins. L'ensemble des agences territoriales de l'ONF concernées par le projet (FD de l'Aigoual, FD de Blois et FD de Mormal) ont également la possibilité d'accéder, sur demande, aux données et métadonnées générées par le projet.

Aucune collecte de données à caractère personnel n'a été envisagée sur la durée du projet et par conséquent le risque relatif aux données à caractère personnel est minimisé.

En matière de protection des données, nous nous sommes conformés aux exigences et à la politique institutionnelle du CNRS.

---

4. Exigences légales et éthiques, codes de conduite

4a. Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?

Aucune collecte de données à caractère personnel n'a été effectuée sur la durée du projet. Par conséquent, le respect des dispositions de la législation sur les données à caractère personnel n'est pas applicable dans le cadre du projet IMPRINT.

4b. Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?

Chaque scientifique impliqué dans le projet a respecté les règles de sa tutelle scientifique, en matière de propriété intellectuelle. Aucun brevet n'est envisagé sur la base du projet IMPRINT. Tous les scientifiques impliqués dans le projet IMPRINT ont été associés aux publications en fonction de leurs contributions respectives (voir la liste des publications associées à ce PGD).

L'ensemble des données déjà déposées sur l'entrepôt InDoRES sont librement accessibles et mise à disposition de tous selon les termes de la Licence Creative Commons Attribution 4.0 International. Une partie des données n'a pas encore été télévérsée sur InDoRES mais nous prévoyons de le faire une fois que les données aurons été validées scientifiquement et valorisées dans le cadre de publications scientifiques.

Le projet IMPRINT étant mono-parteneaire, aucun accord de consortium n'a été signé ou même rédigé avec d'autres partenaires académiques. Cependant, des conventions d'échange et d'utilisation des données ont été passées entre l'UPJV, le CNRS et les différentes agences territoriales de l'ONF, notamment les agences territoriales Hérault-Gard, Val-de-Loire et Nord-Pas-de-Calais.

4c. Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?

Les questions d'éthique n'ont aucune incidence sur la façon dont les données sont stockées et transférées. Les données qui ont été collectées sur la durée du projet n'engage pas d'adopter des codes de conduite nationaux ou internationaux ou de code d'éthique institutionnel.

---

5. Partage des données et conservation à long terme

5a. Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?

A ce stade de la fin administrative du projet, huit articles scientifiques en lien direct avec le projet IMPRINT ont déjà été publiés. L'ensemble des données utilisées dans ces publications scientifiques ainsi que les scripts R rédigés pour analyser les données associées sont disponibles en libre accès sur des plateforme dédiés comme figshare. A ce titre, l'ensemble des données brutes utilisées et des scripts R associées aux 4 articles scientifiques issues de la thèse d'Eva Gril, financée sur le projet IMPRINT, sont disponibles sur le compte figshare d'Eva Gril. D'autre part, une grande partie des données issues des trois FD étudiées dans le cadre du projet IMPRINT, et déjà valorisées au format article scientifique, sont disponibles en accès libre sur l'entrepôt de données InDoRES. L'ensemble des huit articles déjà publiés dans le cadre du projet IMPRINT sont également disponibles au format preprint sur la plateforme HAL.

Les futurs résultats du projet IMPRINT, non encore publié à ce jour, seront, autant que faire se peut, publiés dans des revues à comité de lecture appropriées. Les données utilisées pour la publication des résultats seront également publiées en libre accès au moment de la publication de l'article, ainsi que les scripts (lignes de codes) utilisés pour générer les résultats à partir des données, ceci dans le but d'assurer la reproductibilité des résultats publiés. Une fois valorisées au format article scientifique, les données non encore partagées le seront via l'entrepôts de données InDoRES associé au projet IMPRINT ou via des plateformes de partage de données en libre accès telles que figshare ou dataterra pour les données spatialisées plus volumineuses comme les données brutes des nuages de points LiDAR des trois FD étudiées. A l'issu du processus de publication et en respectant les politiques éditoriales des éditeurs (délais d'embargo spécifiques à chaque revue), les nouvelles publications issues du projet seront déposées dans une archive ouverte et libre d'accès (HAL).

5b. Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?

L'ensemble des données générées par le projet IMPRINT sont vouées à être partagées publiquement. ces données sont uniques et originales et devront donc être conservées sans mention de délais. Aucune donnée ne sera détruite. Les données les plus volumineuses (données LiDAR) seront, dans la mesure du possible, conservées sur le long terme (au-delà de la période du projet) sur des entrepôts de données accessibles à tous, comme l'entrepôt InDoRES sur lequel nous avons déjà déposés une grande partie des données. Ces données, ainsi que l'ensemble des données stockées et archivées dans la géodatabase du projet IMPRINT, seront également conservées en interne au sein du laboratoire EDYSAN, sur un serveur ou un disque dur externe, et disponible sur demande après la publication des résultats.

5c. Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?

Une grande partie des données du projet IMPRINT sont déjà stockées au format tabulaire (.tab), texte (.txt) ou shapefile (.shp) dans l'entrepôt de données InDoRES et donc disponible dans des formats standards et ouverts à des fins de partage et de réutilisation des données. Les données brutes et les scripts d'analyses associées (.R) aux données brutes pour produire de nouvelles données réutilisables (.RData) sont également disponible sur la plateforme figshare et directement réutilisable via R et RStudio.

5d. Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?

Les cinq blocs de données (placettes, microclimats, LiDAR, dendrométrie, flore) déjà téléversée sur lnDoRES dispose d'un identifiant

unique et pérenne (DOI) par bloc (10.48579/PRO/RSBT3C, 10.48579/PRO/ZQBQUW, 10.48579/PRO/XH62TC, 10.48579/PRO/AIXCGM, 10.48579/PRO/BJ9MT6.)

Pour les jeux de données plus volumineux comme les données LiDAR (supérieur à 10 Go), nous envisageons un partage des données avec production d'un DOI via des entrepôt comme dataterra qui sont plus adaptées à ce type de données que l'entrepôt InDoRES.

Enfin, une autre possibilité (payante cette fois-ci) pour les jeux de données LiDAR très volumineux (supérieur à 1 Go) sera de soumettre ces derniers à la revue Scientific Data qui propose également, après une évaluation par les pairs, de passer par une plateforme de stockage parmi une liste de plateformes possibles (p.ex. Dryad, figshare, Zenodo, etc.). Les coûts de publication d'un jeu de données dans cette revue sont de 1570 EUR avec des surcoûts possibles suivant la taille du jeu de données déposé (+775 EUR pour une taille allant de 100 à 250 Go). Cette option payante sera envisagée en dernier recours uniquement pour les jeux de données les plus volumineux et si aucune solution alternative gratuite n'est disponible.

---

6. Responsabilités et ressources en matière de gestion des données

6a. Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?

La mise en oeuvre du plan de gestion des données (PGD) a été réalisée par Jonathan Lenoir (Chargé de Recherche en Ecologie au CNRS), porteur du projet IMPRINT et responsable du PGD et de la gestion globale de ce dernier sur la durée totale du projet (du 1er Octobre 2019 au 31 Mars 2024). Il s'est assuré de sa mise en oeuvre et de sa mise à jour finale après la fin administrative du projet.

La saisie des données, la production des métadonnées, le suivi de la qualité des données ainsi que le stockage et le partage des données et des métadonnées a été assurée par l'ensemble des scientifiques impliqués dans le projet IMPRINT mais plus particulièrement par Emilie Gallet-Moron, Ingénieure d'Etude en géomatique dans l'unité de recherche EDYSAN, et responsable de la création et de la gestion de la géodatabase du projet IMPRINT. En ce qui concerne plus particulièrement les données microclimatiques, comme les séries temporelles de température de l'air et du sol, l'acquisition des données sur le terrain, la mise en forme des données et l'analyse des données a été assurée par la doctorante CNRS, Eva Gril, recrutée sur le projet IMPRINT et sous la supervision de Fabien Spicher (Ingénieur d'Etude en Sciences Forestières et Métrologie) et Jonathan Lenoir (Responsable du projet IMPRINT). Pour ce qui concerne les données d'observations du milieu, comme les relevés dendrométriques et botaniques, les activités de gestion des données ont été également assurées par Eva Gril sous la supervision de Fabien Spicher et Jonathan Lenoir. Les données issues de prélèvements réalisés in-situ et analysés en laboratoire, comme les communautés d'arthropodes du sol, ont été saisies et archivées par Vincent Le Roux (Maître de Conférence en Ecologie des Insectes) et Ronan Marrec (Maître de Conférence en Ecologie du Paysage). Enfin, pour ce qui relève des données cartographiques et issues de la télédéctection, comme les données issues des nuages de points LiDAR, la gestion des données a été assurée par Emilie Gallet-Moron (Ingénieure d'Etude en Géomatique) et Mariane Lasliers (Maître de Conférence en Télédétection).

6b. Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?

En terme de ressources humaines, le recrutement d'Eva Gril, doctorante à 100% sur le projet IMPRINT encadrée par Jonathan Lenoir, Ronan Marrec et Guillaume Decocq (Professeur en Sciences Végétales et Fongiques), a permis d'assurer la collecte mais également la gestion, l'analyse et la valorisation scientifique d'une majeure partie des données générées par le projet IMPRINT. Les données ont également été mises en base de données (création d'une géodatabase) grâce à l'appuie d'une Ingénieure d'Etude en Géomatique dans l'unité EDYSAN, Emilie Gallet-Moron, qui s'est également chargée de formater et de préparer les données au bon format et pour que celles-ci soient bien FAIR (Facile à trouver, Accessible, Interopérable et Réutilisable).

Concernant les frais de stockage, nous avons utilisés au maximum les outils mis à disposition par les tutelles CNRS (sDrive) et UPJV (NextCould) de l'unité EDYSAN pendant toute la durée du projet . Enfin, pour ce qui est du partage des données, nous avons pour l'instant eu recours à l'entrepôt de stockage InDoRES sous la tutelle de l’Institut Ecologie et Environnement du CNRS et du Muséum national d’Histoire naturelle. Pour les jeux de données les plus volumineux (plus de 100 Go à plusieurs To), nous envisageons de recourir à l'entrepôt dataterra.