
DMP du projet "IMPRINT"

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 "IMPRINT"
Langue	fra
Date de création	2020-01-29
Date de dernière modification	2020-03-12
Identifiant	PGD-IMPRINT

Renseignements sur le projet

Titre du projet 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 de l'air 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 fraîches 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 la biodiversité végétale et carabique 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

Produits de recherche :

1. Default research output (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
Jonathan Lenoir - https://orcid.org/0000-0003-0638-9582		<ul style="list-style-type: none"> • Coordinateur du projet • Personne contact pour les données • Responsable du plan

Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

DMP du projet "IMPRINT"

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 qui seront produites dans le cadre du projet IMPRINT sont de quatre types : (i) des **données microclimatiques recueillies in-situ** sur les différents sites d'étude envisagés dans le projet ; (ii) des **données d'observations du milieu réalisées in-situ** (p.ex. données de localisation géographique, relevés dendrométriques et botaniques, etc.) ; (iii) des **données issues de prélèvements réalisés in-situ puis analysés en laboratoire** (p.ex. pièges Barber pour inventorier les communautés d'arthropodes du sol, échantillons de sols, etc.) ; et (iv) des **données issues de la télédétection** (imagerie LiDAR).

(i) Les données microclimatiques (p.ex. T°C et humidité relative) seront recueillies à l'aide de capteurs de données environnementales posés directement in-situ par les scientifiques impliqués dans le projet ou bien posés à distance par l'intermédiaire des agents du Réseau National de suivi à long terme des ECOSystèmes FOREstiers ([RENECOFOR](#)) de l'Office National des Forêts (ONF), auxquels les capteurs seront envoyés par courrier accompagné d'un protocole d'installation in-situ. En effet le projet s'appuie en partie sur le dispositif expérimental créé et géré par le réseau RENECOFOR pour la collecte et la mise à disposition de données existantes. Les capteurs seront posés in-situ autant que faire se peut dès la première année du projet (2020). Plusieurs types de capteurs seront utilisés : (i) des capteurs HOBO UA-001-64 pour les données de T°C de l'air à 1 m au dessus du sol (une fréquence d'un enregistrement toute les heures est visée) ; (ii) des capteurs HOBO UA-001-08 pour les données de T°C à 8 cm sous la surface du sol (une fréquence d'un enregistrement toute les 2 heures est visée) ; et (iii) des capteurs TMS4 qui permettent de prendre la T°C 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 à 8 cm sous la surface du sol (une fréquence d'un enregistrement toute les 15 mins est visée). Les données microclimatiques seront déchargées au moins une fois par an (2021 et 2022), lors des campagnes de terrain, à l'aide des logiciels dédiés [HOBOWare](#) et [Lolly Manager](#) puis analysées sous le logiciel libre [R](#) de traitements statistiques.

(ii) Les données d'observations du milieu (p.ex. localisation des placettes de mesure in-situ, relevés dendrométriques et botaniques, etc.) seront collectées par chaque scientifique impliqué dans le projet. La "*provenance des données*" ou localisation géographique (latitude, longitude, altitude) des placettes de mesures sera documentée dès l'installation des placettes de suivi du microclimat, la première année (2020), à l'aide d'un GPS Trimble et d'une antenne pour une meilleure précision. Les données dendrométriques (e.g. diamètre des arbres à 1,30 m) de structure du peuplement forestier seront mesurées in-situ dès l'installation des placettes en 2020, avec prises de photographies hémisphériques pour quantifier l'ouverture du couvert forestier sur chaque placette. La composition en espèces végétales dans les différentes strates de végétation (muscinale, herbacée, arbustive et arborée) sera inventoriée sur des placettes de surface fixe (400 m²) au cours de la deuxième année du projet (2021). La saisie des données d'observation du milieu se fera sur un tableur Excel à l'aide d'une tablette endurcie et étanche pour faire face aux aléas climatiques liés aux conditions du terrain. L'analyse des données se fera à l'aide du logiciel libre [R](#).

(iii) Les données issues de prélèvements réalisés in-situ et analysés en laboratoire comprennent la pose de pièges Barber pour déterminer la composition des communautés d'arthropodes du sol. Les pièges seront posés in-situ (et récupérés 15 jours plus tard) par l'équipe de scientifique impliqués dans le projet. Deux campagnes de prélèvements sont prévues au printemps et à l'été de l'année 2021, en parallèle des inventaires floristiques (ii). Les données d'identifications taxonomiques seront produites en laboratoire au cours de la troisième année du projet (2022) par l'équipe d'entomologues impliqués dans le projet. L'identification des individus piégés se fera, autant que faire se peut, au rang de l'espèce, du genre ou de la famille, suivant les compétences en interne et d'éventuels prestation de service auprès de spécialistes de certains genres taxonomiques. Les données de composition taxonomique seront saisies dans un tableur Excel avant d'être analysées sous le logiciel libre [R](#).

(iv) Les données issues de la télédétection (images LiDAR) seront collectées par un prestataire externe au projet lors d'un survol des différents sites étudiés. Cette campagne d'acquisition des données LiDAR est prévue au cours de l'année 2021, si possible au printemps ou au début de l'été, lorsque les feuillues commenceront à débourrer et avant la mise en place totale de la canopée haute, ceci dans le but de bien caractériser la structure verticale du sous-étage forestier et d'obtenir un modèle numérique de terrain (MNT) suffisamment précis. Les images LiDAR seront traitées et analysées sous le logiciel [ArcGIS](#) ou sous le logiciel [R](#).

En ce qui concerne les données existantes et qui seront potentiellement 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'Office National des Forêts (ONF) mais aussi des données climatiques (T°C et précipitations) en libre accès (bases de données [WorldClim](#), [CHELSA](#) et [TerraClimate](#)) ou bien issues du réseau Météo France. Nous prévoyons d'établir des conventions d'échanges de données 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 seront 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'agisse 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 électronique (.docx, .pdf).

Les données qui seront acquises au cours du projet sont pratiquement toutes de type **numérique** (tableaux de données et géodatabases). Quelques images (p.ex. photos hémisphériques ou photos traditionnelles) ou vidéos seront également 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.

Les données microclimatiques de T°C et d'humidité relatives seront stockées au format électronique uniquement (.txt, .csv, .xlsx). Les fichiers bruts issus des capteurs HOBO et TMS4 peuvent être enregistrés dans plusieurs formats puis traités par n'importe quel logiciel de traitement de données. Le volume des données microclimatiques qui seront générées pendant la durée du projet est estimé à plusieurs Go. Les données d'observations du milieu comme les données GPS ou bien les données issues des inventaires dendrométriques et floristiques seront enregistrées au format tableur électronique (.txt, .csv, .xlsx). De même, les données d'identification taxonomiques ou bien les données d'analyses de sol (pH, C/N, etc.), acquises en laboratoire, seront enregistrées au format tableur électronique (.txt, .csv, .xlsx). Le volume de ces données est estimé à plusieurs Mo. Les données LiDAR seront stockées au format dédié (.las) à ce type de données 3D ou "point cloud" pour nuage de points 3D, avec un volume de données pouvant dépasser plusieurs To.

Dans la mesure du possible, des formats standards (.txt) et ouverts seront privilégiés à des fins de partage et de réutilisation. Un projet de géodatabase ouverte à l'aide du logiciel [ArcGIS](#) ou de [RShiny](#) sera envisagé en fin de projet afin de mettre en libre accès ou en consultation libre l'ensemble des données acquises dans le cadre du projet IMPRINT.

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 ?

Chaque fichier de données sera accompagné au minimum 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 sera l'anglais pour une plus grande réutilisation potentielle des données collectées. Le standard de métadonnées envisagé dans le cadre du projet est le standard "*Ecological Metadata Language*" ([EML](#)) avec une implémentation possible sous le logiciel [R](#) grâce au [package EML](#).

La convention de nommage des fichiers ou dossiers de données sera 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 précisons la version dans le nom du fichier (p.ex. "20200218_DMP_IMPRINT_v01.docx").

Une documentation détaillée des données utilisées ainsi que des analyses effectuées à partir des données sera également produite au format [RMarkdown](#) (.rmd) à partir du 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 seront suivies 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 prévoit

suffisamment de répliqués. Au total, trois Forêts Domaniales à dominante feuillue seront 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) seront (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 prévoyons également d'équiper une partie du réseau RENECOFOR (les placettes de forêts feuillues notamment) avec un jeu apparié de sondes TMS4 : (i) une sonde sous-couvert forestier au sein de la placette de suivie du réseau RENECOFOR et (ii) une autre sonde 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 sera très complémentaire du précédent qui est focalisé sur trois forêts domaniales équipées d'une plus grande densité de capteurs microclimatiques.

Concernant la qualité et la conformité de la collecte des données microclimatiques, il est prévu une phase d'intercalibration des capteurs HOBO UA-001-08, HOBO UA-001-64 et TMS4 en conditions contrôlées. L'installation des capteurs de T°C et d'humidité relative du sol, in-situ, suivra 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 seront 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 (du 1er Octobre 2019 au 30 Septembre 2023), les données et les métadonnées générées seront hébergées sur les différents postes de travail chiffrés des membres du projet IMPRINT. Pour les données les plus volumineuses (données LiDAR), il est prévu l'achat d'un disque dur externe, qui sera également chiffré. Pour les données les moins volumineuses (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, il est prévu d'utiliser un dossier partagé et intitulé "IMPRINT" sur la plateforme de sauvegarde et de partage sécurisé du CNRS : [myCoRe](#). Ce dossier partagé entre les scientifiques impliqués dans le projet a déjà été créé.

Des sauvegardes (backup) régulières (tous les 6 mois au moins) de l'ensemble des données et métadonnées du projet (inclues les données LiDAR) seront également effectuées et enregistrées sur le serveur NAS de l'unité de recherche [EDYSAN](#) (UMR 7058 CNRS). A noter également que les données LiDAR seront partagées avec chaque agence territoriale de l'ONF concernée par le projet (FD de l'Aigoual, FD de Blois et FD de Mormal). Cela assurera une sauvegarde supplémentaire.

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 ([myCoRe](#)). 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 envisageons également l'enregistrement des données LiDAR sur un cloud avec un espace de stockage conséquent et en accord avec la plateforme [MATRICS](#) de l'Université de Picardie Jules Verne (UPJV).

L'ensemble des scientifiques impliqués dans le projet aura 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 se fera 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) auront également accès aux données et métadonnées générées par le projet.

Aucune collecte de données à caractère personnel n'est 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 conformeront 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'est envisagée 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 respectera 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 seront associés aux publications en fonction de leurs contributions respectives.

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

Les questions d'éthique n'auront aucune incidence sur la façon dont les données seront stockées et transférées. Les données qui seront collectées sur la durée du projet n'engagent 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 ?

Les résultats obtenus au cours du projet 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. Les données seront partagées via des plateformes de partage de données en libre accès telles que [figshare](#). A l'issue du processus de publication et en respectant les politiques éditoriales des éditeurs (délais d'embargo spécifiques à chaque revue), les 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 seront uniques 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. Les données LiDAR seront 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 ?

Dans la mesure du possible, des formats standards et ouverts seront privilégiés à des fins de partage et de réutilisation des données.

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

Pour les jeux de données peu à moyennement volumineux (maximum 1 Go), nous privilégierons l'utilisation de la plateforme [figshare](#) permettant le partage gratuit des données avec l'attribution d'un identifiant unique et pérenne (DOI). A titre d'exemple, le porteur du projet IMPRINT a récemment publié un article en premier auteur dans lequel les données et le code utilisés pour générer les résultats présentés dans la publication sont disponibles en libre accès via [figshare](#), avec un DOI associé : doi.org/10.6084/m9.figshare.10012841

Pour les jeux de données plus volumineux comme les données LiDAR (supérieur à 10 Go), nous envisagerons un partage des données via des plateformes comme [Zenodo \(max 50 Go\)](#) avec possibilité de contacter la plateforme pour plus de capacité de stockage si nécessaire) ou bien le Centre informatique national de l'enseignement supérieur ([Cines](#)) dont le CNRS est [partenaire](#).

Enfin, une autre possibilité ([payante cette fois-ci](#)) pour les jeux de données 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 pour les jeux de données les plus volumineux.

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. Il endossera la responsabilité du PGD et la gestion globale de ce dernier sur la durée totale du projet (du 1er Octobre 2019 au 30 Septembre 2023). Il s'assurera de la mise en oeuvre et de la mise à jour régulière du PGD sur la durée 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 sera assurée par l'ensemble des scientifiques impliqués dans le projet IMPRINT. En ce qui concerne les données microclimatiques, comme la température, l'ensemble des activités mentionnées dans la phrase précédente sera assurée par la doctorante recrutée (Eva Gril) sur le projet IMPRINT et sous la supervision de Fabien Spicher (Ingénieur d'Etude en Sciences Forestières et Métrologie) et Jonathan Lenoir. Pour ce qui concerne les données d'observations du milieu, comme les relevés dendrométriques et botaniques, les activités de gestion des données seront 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, seront 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 télédétection, comme les images LiDAR, la gestion des données sera assurée par Emilie Gallet-Moron (Ingénieure d'Etude en Géomatique) et Jonathan Lenoir.

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), permettra d'assurer que les données générées par le projet IMPRINT seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable).

Concernant les frais de stockage et de partage des données, nous utiliserons au maximum les plateformes de stockage gratuites, notamment pour les jeux de données de taille moyenne. Pour les jeux de données les plus volumineux (plus de 100 Go à plusieurs To), il sera éventuellement envisagé de recourir à des plateformes payantes pour la mise en libre accès (cf. Scientific Data). Pour cela, une partie du budget IMPRINT sera potentiellement alloué à la publication en libre accès (max 2000 EUR). Suivant les ressources financières disponibles et les besoins en matière de publication des données en libre accès, il sera également envisagé de faire une demande de soutien financier via l'action 3 de l'appel S2R ([Soutien au Rayonnement de la Recherche](#)) de l'UPJV. Cette action permet d'apporter une aide financière aux chercheurs qui souhaitent publier un article ou un ouvrage scientifique avec comme condition que l'unité de recherche concernée (EDYSAN dans ce cas) se doit de participer financièrement à la publication. Dans ce cas précis, la participation financière de l'unité EDYSAN pourra provenir d'une partie du budget associé au projet IMPRINT.

