

# Rapport final de vérification des données prises lors de l'installation des sentiers à la CEB, Gabon

Gauthier Ligot et Alexandre Ernst

01/09/2020

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Description de l'aquisition des données</b>	<b>1</b>
2.1	Enregistrement des données . . . . .	1
2.2	Description des attributs . . . . .	2
2.3	Calendrier de l'acquisition des données . . . . .	5
2.4	Problèmes techniques . . . . .	5
<b>3</b>	<b>Présentation des données</b>	<b>5</b>
3.1	Bloc sud . . . . .	5
3.2	Bloc nord . . . . .	12
<b>4</b>	<b>Conclusion</b>	<b>17</b>

## 1 Introduction

Ce document présente les données récoltées lors de l'installation des sentiers de suivi de la croissance d'essences tropicales dans la concession Precious au Gabon. Il a pour but de présenter en détail les données récoltées, la méthode d'encodage des données, leur archivage et les problèmes techniques rencontrés. Les données sont également présentées pour identifier l'importance des données manquantes et erronées.

L'objectif de ce document, un peu technique, est que toutes personnes n'ayant pas participé au projet puisse, prendre en main la gestion et l'analyse des données.

## 2 Description de l'aquisition des données

### 2.1 Enregistrement des données

Depuis le début de l'installation, les données ont toujours été encodées sur le terrain à l'aide d'un smartphone en utilisant l'application Open Foris Collect Mobile. En fin de semaine de travail (les samedis), les données étaient exportées du téléphone sur un Google drive (Login : natureplus.gsm). Les sauvegardes (fichier.collect-data) étaient ensuite rapidement parcourues par Gauthier Ligot ou Franck Monthe avant d'être placées dans le dossier OF\_mobile/données\_backup du Google drive. La vérification des données à ce stade était très sommaire : présence de données et de photos et vérification du nombre d'arbres mesurés par jour de travail. Néanmoins, le formulaire développé sur Open Foris Collect permettait déjà un contrôle des données sur le terrain. Les données présentes sur le Google drive (y compris l'ensemble des sauvegardes intermédiaires) ont ensuite été archivées sur un disque dur externe de Gauthier Ligot.

L'ensemble des sauvegardes (fichiers.collect-data) a été chargé dans un projet Open Foris Collect sur un PC pour ensuite extraire deux fichiers Excel : cluster.xlsx et tree.xlsx. Ces deux fichiers contiennent l'ensemble des données récoltées à l'exclusion des photographies. Ces données sont également archivées sur le dépôt GitHub. Elle seront prochainement ajoutées à la base de données SQLITE des sentiers DynAfFor-P3FAC.

## 2.2 Description des attributs

### 2.2.1 Fichier cluster.xlsx

Les données ont été enregistrées par cluster. Généralement un cluster contient les données collectées par une même équipe de terrain, dans un même bloc et pendant une seule journée. Le fichier cluster.xlsx contient les caractéristiques de chaque cluster. On retrouve dans ce fichier les colonnes suivantes :

- **date\_year** : Année du cluster.
- **date\_month** : Mois du cluster.
- **date\_day** : Jour du cluster.
- **bloc** : NORD ou SUD. Indique dans quel bloc le cluster a été mesuré.
- **operator** : Nom du responsable du cluster.
- **trail\_center\_srs** : Système de coordonnées utilisé.
- **trail\_center\_x** : Coordonnée x du sentier (centre du bloc)
- **trail\_center\_y** : Coordonnée y du sentier (centre du bloc)
- **fm\_u** : Nom de l'UFA
- **aac** : Nom de l'assiette annuelle de coupe

### 2.2.2 Fichier tree.xlsx

Le fichier tree.xlsx contient l'ensemble des observations récoltées pour chaque arbres. On retrouve dans ce fichier les colonnes suivantes :

- **cluster\_date\_year** : Année du cluster.
- **cluster\_date\_month** : Mois du cluster.
- **cluster\_date\_day** : Jour du cluster.
- **cluster\_bloc** : NORD ou SUD. Indique dans quel bloc le cluster a été mesuré.
- **tag** : Numéro de la plaquette du préinventaire.
- **stem\_letter** : Lettre identifiant la tige mesurée (prévue pour mesurer les arbres multitiges).
- **number** : Numéro peint sur l'arbre pendant l'installation des sentiers
- **species** : code de l'espèce de l'arbre.
- **diameter\_preinventory** : Diamètre mesuré lors du préinventaire (en cm).
- **date\_year** : Année de la mesure.
- **date\_month** : Mois de la mesure.
- **date\_day** : Jour de la mesure.
- **diameter** : Diamètre mesuré lors de l'installation du sentier (en cm).
- **diameter\_unit\_name** : colonne Open-Foris pas utilisée.
- **hom** : Hauteur du niveau de mesure (en cm).
- **hom\_unit\_name** : Colonne Open-Foris pas utilisée.
- **code\_observation[1]** : Premier code éventuel d'une observation codifiée.
- **code\_observation[2]** : Deuxième code éventuel d'une observation codifiée.
- **code\_observation[3]** : Troisième code éventuel d'une observation codifiée.
- **code\_observation[4]** : Quatrième code éventuel d'une observation codifiée.
- **code\_observation[5]** : Cinquième code éventuel d'une observation codifiée.
- **habitat** : Code décrivant le type de forêt de l'arbre mesuré.
- **coordinates\_srs** : Système de coordonnées utilisé lors du préinventaire.
- **coordinates\_x** : Coordonnée x de l'arbre utilisé lors du préinventaire.
- **coordinates\_y** : Coordonnée y de l'arbre utilisé lors du préinventaire.
- **coordinates\_new\_srs** : Système de coordonnées utilisé lors de l'installation des sentiers.
- **coordinates\_new\_x** : Coordonnée x de l'arbre utilisé lors de l'installation des sentiers.

TABLE 1 – Description des codes utilisés pour les observations.

code	nom
0	normal
1	asymetrique
2	liane
3	termite
4	etrangleur
5	pourriture
6	mourrant
7	etete
8	blessure
9	incline
10	relascope
11	exploitation
12	rejet
13	hauteur
97	disparu
98	missingData
99	autre
100	erreur

- **coordinates\_new\_y** : Coordonnée y de l'arbre utilisé lors de l'installation des sentiers.
- **dawkins** : Code de dawkins.
- **log\_length** : Code de qualité du bois (longueur de la grume).
- **photo** : Nom de la photo éventuellement prise.
- **comment** : Commentaire supplémentaire.

### 2.2.3 Dictionnaires

La description des codes utilisés est présentée dans les tableaux 1, 2, 3 et 4. Ces tableaux sont également sauvegardés sous forme de fichiers texte (.csv) sur le dépôt GitHub.

TABLE 2 – Description des codes de mortalités utilisés.

code	nom
MSP	Mort sur pied naturelle
CHA1	Chablis primaire naturelle
CHA2	Chablis secondaire naturelle
EXP	Arbre exploité
MSPE	Arbre mort sur pied suite aux blessures occasionnées par l'exploitation (y compris débardage)
EXPV	Arbre abattu par des villageois (bois, miel, etc.)
MSPV	Arbre écorcé ou brûlé
CHAV	Arbre entraîné par la chute d'un arbre exploité par des villageois
INCO	Arbre mort d'origine inconnue

TABLE 3 – Description des codes de dawkins utilisés.

code	nom	description
1	sous-bois partie basse	Arbre entièrement ombragé verticalement et latéralement par les houppiers d'autres arbres
2	sous-bois partie haute	Arbre entièrement ombragé verticalement mais bénéficiant de lumière latéralement
3	canopée partie basse	Arbre partiellement ombragé verticalement
4	canopée	Arbre non ombragé verticalement mais ombragé latéralement
5	émergent	Arbre non ombragé verticalement et latéralement (pas d'autre houppier dans un cône inversé de 90° dans lequel se trouve l'arbre cible)

TABLE 4 – Description des codes de dawkins utilisés.

code	description
H > 15 m	hauteur du fût bien conformé > 15 m
H > 9 m	hauteur du fût bien conformé > 9 m
H > 6 m	hauteur du fût bien conformé > 6 m
H < 6 m	hauteur du fût bien conformé < 6 m

## 2.3 Calendrier de l'acquisition des données

L'installation des sentiers a débuté en septembre 2019 lors d'une mission de Jean-François Gillet. Les premières mesures ont été prises du 16 au 18 septembre dans le bloc nord. Il a ensuite été décidé de continuer les mesures dans le bloc sud. Des mesures ont été réalisées dans le bloc sud du 28 septembre au 13 mars. L'équipe a ensuite effectué des mesures dans le bloc nord du 13 mars au 17 juillet avant de compléter le jeu de données par quelques mesures dans le bloc sud du 17 juillet au 30 juillet.

## 2.4 Problèmes techniques

Plusieurs versions du formulaire Open Foris Collect Mobile ont été utilisées (les versions numérotées 15, 16 et 19). Dans les deux premières versions utilisées du formulaire (15 et 16), la date des enregistrements ne se sauvegardait pas correctement. Lors de chaque ouverture de l'application la date de l'enregistrement s'actualisait et écrasait ainsi les dates de mesures prises sur le terrain. Afin d'installer le correctif (version 19) sur les smartphones, les mesures de terrain ont été suspendues pendant une semaine (semaine 47). Pour les enregistrements encodés avec les versions 15 et 16, les dates de mesures ont été corrigées manuellement. Pour chaque semaine de travail (et donc sauvegarde), une date unique a été fixée pour tous les enregistrements correspondants.

A ce jour, il n'est pas encore possible de réimporter les photos prises avec les versions 15 et 16 dans les projets Open Foris Collect utilisant la version 19. Les photos ne sont néanmoins pas perdues.

Par contre, pour certaines sauvegardes (26/10/2019, 2/11/2019, 10/11/2019, 16/11/2019 et 2/12/2019), les photos n'ont pas été exportées depuis le smartphone vers le Google drive (l'option "exclude binary files" avait été erronément cochée).

Au fur et à mesure de l'avancement du projet, les données s'accumulaient dans le smartphone de terrain. Cette situation s'est avérée ne pas être efficace pour plusieurs raisons : l'équipe de terrain pouvait par erreur modifier des données antérieures et l'accumulation des données remplissait la mémoire des téléphones (en raison de la présence de photos) et ralentissait très fortement la vitesse des sauvegardes vers le Google drive. Pour pallier à ce problème, à partir du 20 janvier 2020, il a été décidé qu'une fois les données sauvegardées et vérifiées sur le Google drive, ces données étaient supprimées des smartphones.

Lors des premières mesures, les coordonnées GPS des arbres n'ont pas été reprises à l'aide du smartphone. Le problème a été réglé par la suite et les coordonnées GPS des arbres suivant ont été relevées.

## 3 Présentation des données

### 3.1 Bloc sud

Au total, le bloc sud contenait 1735 arbres à mesurer. 1686 ont effectivement été trouvés et 1505 arbres ont fait l'objet d'une mesure de diamètre. Les données manquantes sont décrites plus amplement dans les section suivantes.

#### 3.1.1 Arbres non retrouvés

Il est difficile de déterminer les raisons pour lesquelles 54 arbres n'ont fait l'objet d'aucune mesure et d'aucun commentaire. On peut présumer qu'il s'agit essentiellement d'arbres non retrouvés. La figure 1 illustre la répartition spatiale de ces arbres. Ces arbres sont dispersés de manière homogène sur la carte. Nous avons choisi de ne pas faire une journée de terrain pour les retrouver. Ces 54 arbres sont donc abandonnés.

#### 3.1.2 Arbres retrouvés mais abandonnés

Parmi les arbres encodés, il y a 181 enregistrements pour lesquels le diamètre n'a pas été encodé. Pour chacun d'entre eux au moins une observation a été indiquée (ex. un défaut). Pour 172 d'entre eux un commentaire a été indiqué et pour 87 d'entre eux une photo a été prise. A partir de ces informations, on peut donc analyser les raisons de l'abandon des mesures du diamètre.

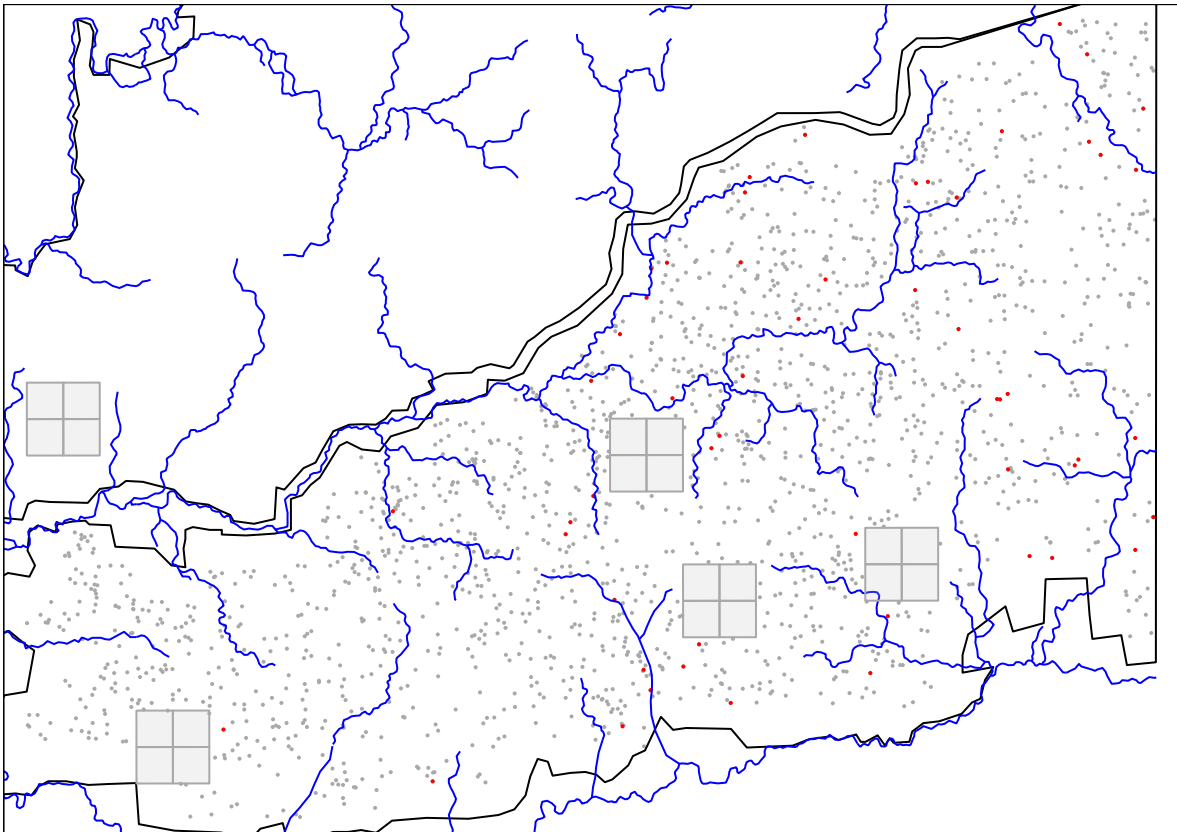


FIGURE 1 – Répartition spatiale des arbres mesurés (en gris) et des arbres non retrouvés (en rouge). Les limites des blocs sont en noires et les cours d'eau en bleu.

TABLE 5 – Liste et occurrence des commentaires inscrits pour les enregistrements sans mesure du diamètre.

Catégories de commentaires	n
Confusion d'essence	64
Déformation importante	60
Mort ou mourrant	17
NA	16
Introuvable	11
Situé dans une parcelle	8
Liane	3
Blessure	1
Danger	1

TABLE 6 – Les enregistrements avec des doublons du numéro de l'arbre

Tag	Numéro	Essence	Pré-inventaire	Diamètre
7059	160	niove	44	43.7
7059	160	niove	44	43.0
171422	8	andok	45	28.8
171422	8	andok	45	22.1
19244	47	okoume	43	40.3
19244	47	okoume	43	21.8

En grande majorité, il s'agit d'arbres présentant des déformations importantes (Figure 2) ou dont l'essence n'avait pas été identifiée correctement lors du pré-inventaire (Tableau 5).

### 3.1.3 Arbres retrouvés et mesurés

**3.1.3.1 Doublons** Parmi les 1505 arbres mesurés, on trouve 6 arbres pour lesquels le numéro a été utilisé plusieurs fois (Tableau 6). Notons, que si l'on ajoute les relevés sans mesures de diamètre, le nombre de doublons augmente. En analysant la date et la position des différents relevés, il est parfois possible d'identifier quel relevé est erroné. Mais ce n'est pas toujours le cas et jamais avec une grande certitude. Ces doublons seront vérifiés lors du prochain remesurage.

**3.1.3.2 Mesures du diamètre** Pour les 1505 mesures de diamètre, il ne semble pas y avoir de valeur aberrante. Le diamètre minimum est de 9.6 cm et le diamètre maximum est de 174.2 cm. Si l'on compare les valeurs mesurées avec celles du préinventaire, alors on peut néanmoins observer des différences assez importantes. La différence entre les deux mesures excède 10 ou 20 cm dans 187 et 54 cas respectivement.

**3.1.3.3 Hauteur de mesure** Parmi les 1505 mesures de diamètre, on identifie aucune valeur aberrante de la hauteur de mesure. La hauteur de mesure minimum est de 130 cm et la hauteur de mesure maximum est de 450 cm.

**3.1.3.4 Type de forêt** Le tableau 7 indique le nombre d'observations par type de forêt.

**3.1.3.5 Indice de dawkins** Le tableau 8 indique la répartition des arbres mesurés par classe de dawkins. Il n'y a pas de valeur aberrante.

**3.1.3.6 Longueur de grume** Le tableau 9 indique la répartition des arbres mesurés par classe de longueur de grume. Il n'y a pas de valeur aberrante.

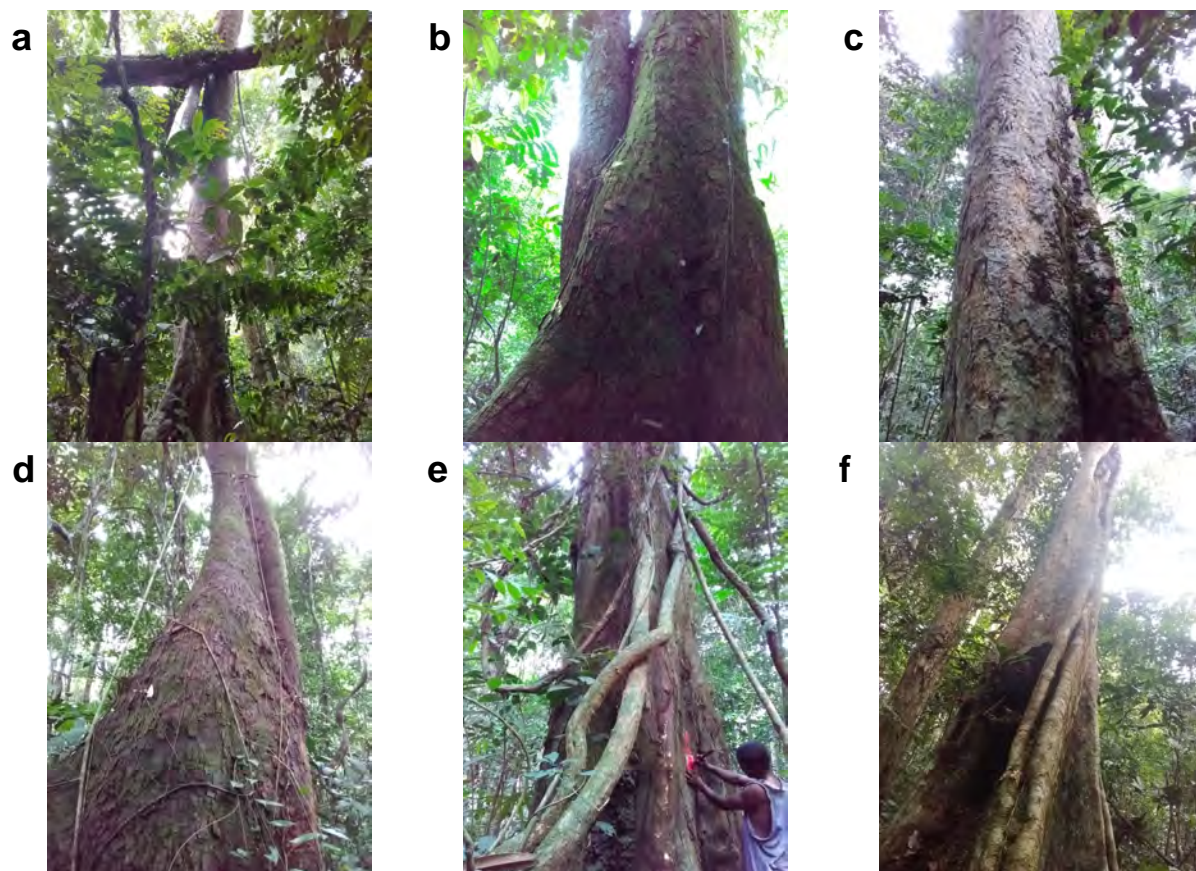


FIGURE 2 – Exemples de photos d’arbres abandonnés en raison d’un danger important (a), de déformations importantes ou de présence de lianes (e).

TABLE 7 – Nombre d’observations par type de forêt.

Type de forêt	n
Vieille forêt secondaire	1435
Forêt marécageuse	56
Forêt périodiquement inondée	10
Forêt à Marantaceae	2
Forêt colonisatrice	1
Jeune forêt secondaire	1

TABLE 8 – Nombre d’arbres mesurés par classe de dawkins.

code	Dawkins	n
1	sous-bois partie basse	77
2	sous-bois partie haute	276
3	canopée partie basse	302
4	canopée	534
5	émergent	316



TABLE 9 – Nombre d’arbres mesurés par classe de longueur de grume.

code	Grume	n
1	H > 15 m	213
2	H > 9 m	498
3	H > 6 m	437
4	H < 6 m	357

TABLE 10 – Occurences des différentes observations.

observation_label	n
NA	1011
normal	458
asymetrique	246
blessure	103
liane	102
etete	93
etrangleur	54
exploitation	48
mourrant	32
pourriture	24
termite	8

**3.1.3.7 Observations** Pour chaque arbre, il était possible d’identifier plusieurs remarques. Il y a donc plus de remarques que d’arbres mesurés. La figure 10 illustre l’importance des différentes observations.

**3.1.3.8 Liste des commentaires** L’ensemble des commentaires écrits pour les arbres avec une mesure de diamètre sont repris dans le tableau 11. On note que les espèces du genre *Dialium* ont régulièrement fait l’objet de remarques (plus de 40 fois).

**3.1.3.9 Effectifs mesurés par essence et par classe de diamètre** La majorité des arbres prévus dans le plan d’échantillonnage ont bien été mesurés. La figure 3 permet de visualiser rapidement la répartition des effectifs par classe de diamètre et par essence. On note qu’il manque essentiellement des arbres de grandes dimensions pour l’andok, le beli, le movingui et le padouk.

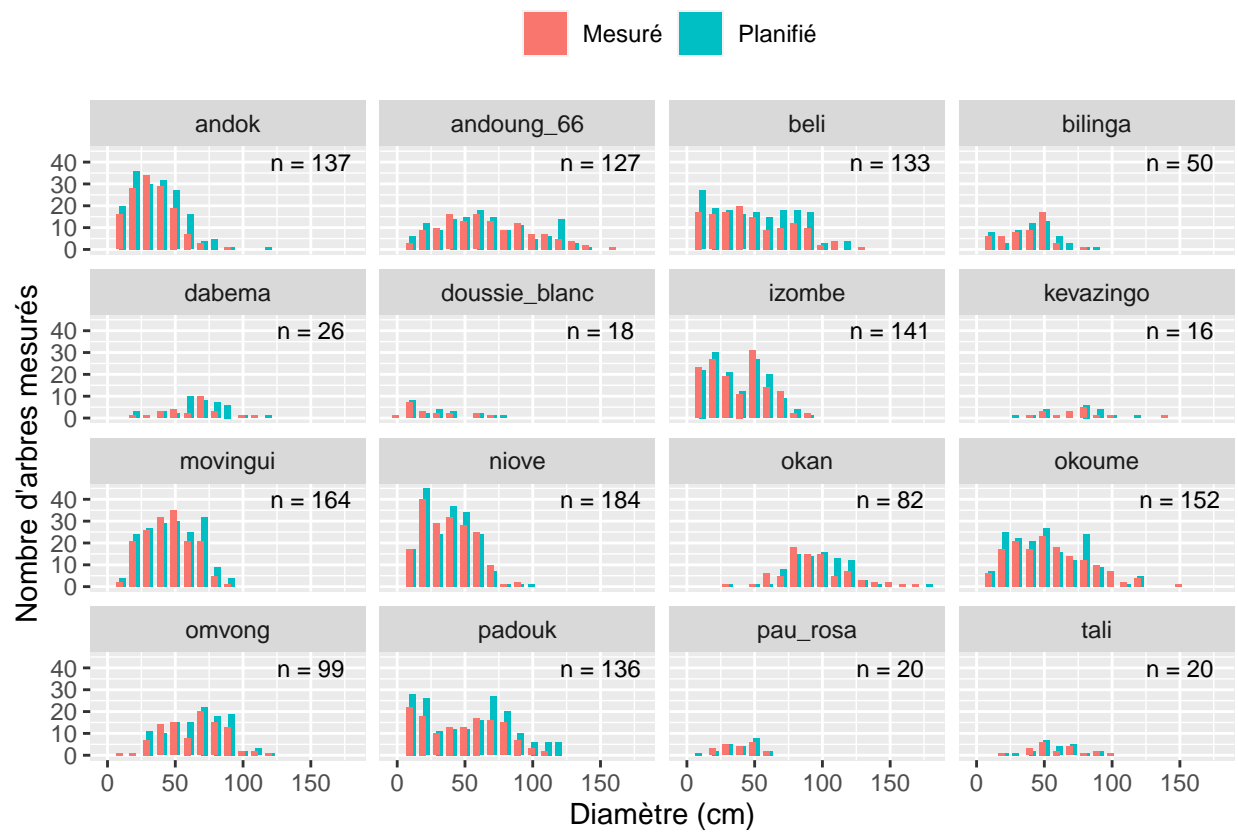


FIGURE 3 – Répartition des effectifs mesurés par essence et classe de diamètre.

TABLE 11 – Liste des commentaires.

comment	n
NA	1441
Dialium pachiphilum	29
Dialium cf lopense	16
Au bord de la rivière	3
Bord de la rivière	2
Il pourrait être un faux padouk	2
Arbre tordu avec des contreforts jusqu'au trait de mesure et une grosse bosse à partir de 250 cm en montant	1
Avec rejet sur le tronc	1
Chablis	1
Chablis, Quelques branches cassées par un autre arbre tombé	1
Déjà mort	1
Déjà tombé et mort	1
Dialium cf lepense	1
Diamètre de moins de 10 cm à 130 cm	1
Guibourtia cf Pellegriniana	1
Guibourtia tessmanii	1
Mourant	1
Tordu	1

## 3.2 Bloc nord

Au total, le bloc nord contenait 1828 arbres à mesurer. 1784 ont effectivement été trouvés et 1627 arbres ont fait l'objet d'une mesure de diamètre.

### 3.2.1 Arbres non retrouvés

Il est difficile de déterminer les raisons pour lesquelles 48 arbres n'ont fait l'objet d'aucune mesure et d'aucun commentaire. On peut présumer qu'il s'agit essentiellement d'arbres non retrouvés. La figure 4 illustre la répartition spatiale de ces arbres. Ces arbres sont dispersés de manière homogène sur la carte. Nous avons choisi de ne pas faire une journée de terrain pour les retrouver. Ces 48 arbres sont donc abandonnés.

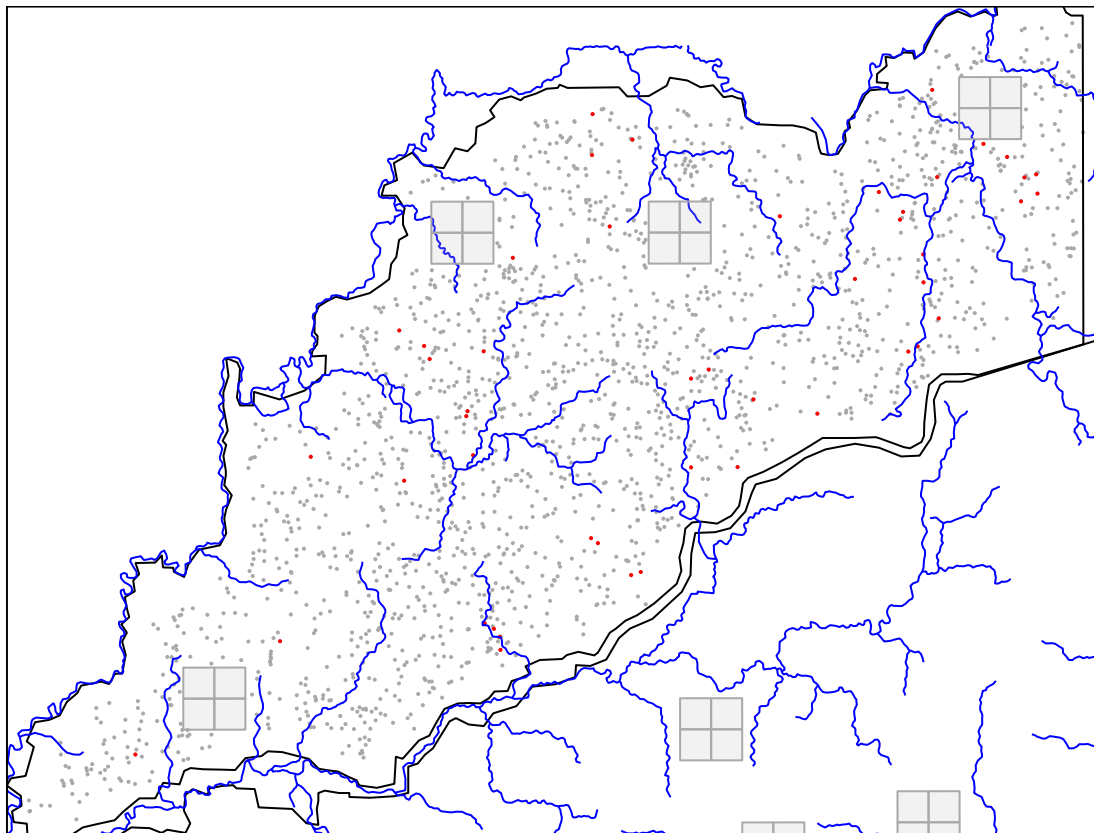


FIGURE 4 – Répartition spatiale des arbres mesurés (en gris) et des arbres à mesurer (en rouge). Les limites des blocs sont en noires et les cours d'eau en bleu.

### 3.2.2 Arbres retrouvés mais abandonnés

Parmi les arbres encodés, il y a 157 enregistrements pour lesquels le diamètre n'a pas été encodé. Pour 155 d'entre eux un commentaire a été indiqué (Tableau 12), pour 157 au moins une observation a été indiquée et pour 71 d'entre eux une photo a été prise.

### 3.2.3 Arbres retrouvés et mesurés

**3.2.3.1 Doublons** Parmi les 1627 arbres mesurés, on trouve 4 doublons du numéro des arbres (13). Un arbre n'a pas été mesuré et en remarque l'équipe a écrit "déjà fait" (izombe 93).

TABLE 12 – Les commentaires inscrits pour les enregistrements sans mesure du diamètre

Catégories de commentaires	n
Déformation importante	67
Confusion d'essence	53
Situé dans une parcelle	18
Mort ou mourrant	13
Introuvable	2
NA	2
Déjà mesuré	1
Liane	1

TABLE 13 – Les enregistrements avec des doublons du numéro de l'arbre

Tag	Numéro	Essence	Pré-inventaire	Diamètre
162306	22	doussie_blanc	16	16.3
162306	22	doussie_blanc	16	14.3
571	129	niove	63	61.4
571	129	niove	63	10.5

Parmi les arbres encodés, il y a 157 enregistrements pour lesquels le diamètre n'a pas été encodé. Pour chacun d'entre eux au moins une observation a été indiquée (ex. un défaut). Pour 155 d'entre eux un commentaire a été indiqué et pour 71 d'entre eux une photo a été prise. A partir de ces informations, on peut donc analyser les raisons de l'abandon des mesures du diamètre.

**3.2.3.2 Mesures du diamètre** Parmi les 1627 mesures de diamètre, il ne semble y avoir aucune valeur aberrante. Le diamètre minimum est de 10 cm et le diamètre maximum est de 168.3 cm. Si l'on compare les valeurs mesurées avec celles du préinventaire, alors on peut néanmoins observer des différences assez importantes. La différence entre les deux mesures excède 10 et 20 cm dans 193 et 53 cas respectivement.

**3.2.3.3 Hauteur de mesure** Parmi les 1627 mesures de diamètre, on identifie aucune valeur aberrante de la hauteur de mesure. La hauteur de mesure minimum est de 110 cm et la hauteur de mesure maximum est de 450 cm.

**3.2.3.4 Type de forêt** Le tableau 14 indique le nombre d'observations par type de forêt.

**3.2.3.5 Indice de dawkins** Le tableau 15 indique la répartition des arbres mesurés par classe de dawkins. Il n'y a pas de valeur aberrante.

TABLE 14 – Nombre d'observations par type de forêt.

Type de forêt	n
Vieille forêt secondaire	1589
Forêt marécageuse	35
Forêt dense ombrophile semi-sempervirente	2
Jeune forêt secondaire	1

TABLE 15 – Nombre d’arbres mesurés par classe de dawkins.

code	Dawkins	n
1	sous-bois partie basse	63
2	sous-bois partie haute	286
3	canopée partie basse	403
4	canopée	499
5	émergent	376

TABLE 16 – Nombre d’arbres mesurés par classe de longueur de grume.

code	Grume	n
1	H > 15 m	108
2	H > 9 m	355
3	H > 6 m	555
4	H < 6 m	609

**3.2.3.6 Longueur de grume** Le tableau 16 indique la répartition des arbres mesurés par classe de longueur de grume. Il n’y a pas de valeur aberrante.

**3.2.3.7 Observations** Pour chaque arbre, il était possible d’identifier plusieurs remarques. Il y a donc plus de remarques que d’arbres mesurés. La figure 17 illustre l’importance des différentes observations.

**3.2.3.8 Liste des commentaires** L’ensemble des commentaires encodés pour les arbres avec une mesure de diamètre est repris dans le tableau 18.

**3.2.3.9 Effectifs mesurés par essence et par classe de diamètre** La majorité des arbres prévus dans le plan d’échantillonnage ont bien été mesurés. La figure 5 permet de visualiser rapidement la répartition des effectifs par classe de diamètre et par essence.

TABLE 17 – Occurences des différentes observations.

observation_label	n
NA	1266
normal	312
blessure	111
asymetrique	78
etete	57
etrangleur	47
mourrant	40
liane	21
pourriture	9
termite	7
exploitation	5

TABLE 18 – Liste des Commentaires.

comment	n
NA	1463
Dialium pachyphyllum	80
Dialium cf lopense	45
Guibourtia cf pellegriniana	10
Guibourtia tesmanni	10
172668 sur la plaquette et 172468 sur la carte	1
172671 sur la plaquette et 172471 sur la carte	1
Branches cassées	1
Cannele	1
Couché	1
Déclassé, contreforts à plus de 4m50	1
Defaut a 130,	1
Encore vivant	1
Erreur diam preinv	1
Hauteur de mesure sur contrefort,	1
La deuxième échelle est déjà cassée	1
La plaquette de ce padouk était portée par l'Ozigo juste à côté et le numéro que portait le padouk est le 168076	1
La plaquette de cet izombé était porté par l'Ozigo juste à côté et le numéro que portait l'Izombé est le 168078	1
Le bon emplacement mais sur la plaquette, 172676 au lieu de 172476, On l'a pris,	1
Le bon emplacement mais sur la plaquette, 172680 au lieu de 172480, On l'a pris,	1
Leger defaut	1
Plus de plaquette,	1
Prise de mesure entre contrefort et bosse,	1
Tordu	1

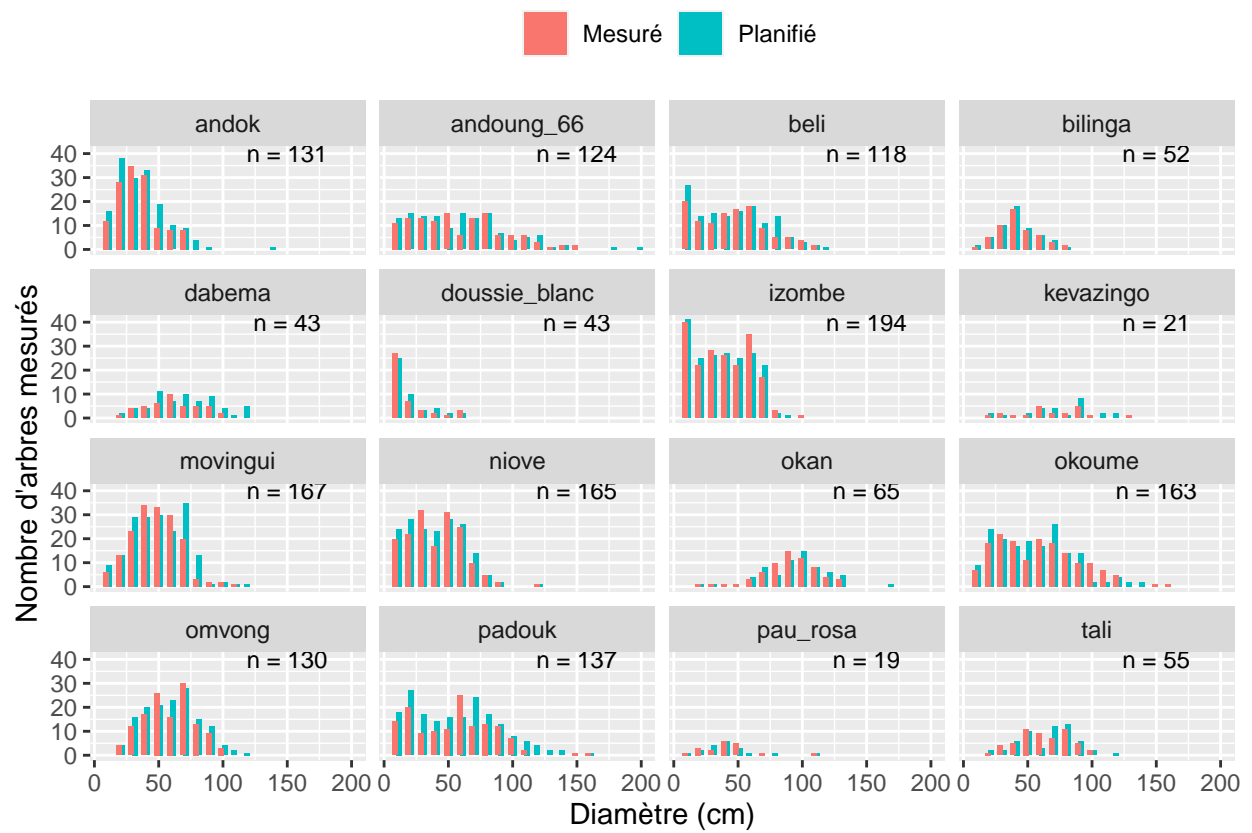


FIGURE 5 – Répartition des effectifs mesurés par essence et classe de diamètre.



## 4 Conclusion

La majorité des arbres a bien été mesurée (3132 arbres mesurés sur 3563 arbres prévus soit 88%). Seulement 54 arbres pour le bloc sud et 48 arbres pour le bloc nord n'ont pas été remesurés et vraisemblablement pas retrouvés. La méthode d'encodage a donc été globalement appropriée. L'échantillonnage obtenu semble suffisant et il n'est pas prévu d'identifier de nouveaux arbres à mesurer.