

# Master 2 : Science et Technologie du Vivant et de l'Environnement (STVE)

## Spécialité Ecologie Biodiversité Evolution (EBE)



*Parcours Ecologie Théorique et Modélisation*

Modélisation de la mortalité des arbres en forêt tropicale  
humide : Dispositif permanent d'Uppangala dans les Ghâts  
occidentaux d'Inde

Jimmy Le Bec

Année universitaire 2010 - 2011

Responsables : Raphaël Pélissier  
Benoît Courbaud

Durée du stage : 15 février – 15 juin 2010





## Remerciements

Ce stage fut court mais néanmoins riche et m'a offert l'occasion de travailler avec de nombreuses personnes. Je tiens donc tout d'abord à remercier l'équipe de l'Institut Français de Pondichéry avec qui j'ai pu réaliser une mission de deux semaines sur le dispositif d'Uppangala. Je remercie ensuite les membres de l'équipe III du laboratoire AMAP à Montpellier qui ont pu m'aider sur des points techniques ou théoriques de ce stage. Je remercie tout patriculièrement Raphael Pélissier qui a encadré ce stage ainsi que Gilles Le Moguedec pour son soutien en statistique. Je remercie enfin Benoît Courbaud du CEMAGREF pour ses conseils et son appui.

# Sommaire

---

Remerciements .....	3
1. Introduction.....	6
1.1. Enjeux de la compréhension de la mortalité.....	6
1.2. Processus biologiques et mortalité en forêt .....	7
1.2.1. Compétition.....	7
1.2.2. Sénescence .....	8
1.2.3. Chute d'arbre : le chablis.....	8
1.3. Modélisation de la mortalité en forêt .....	9
1.3.1. 1.4.1. Prise en compte des différentes covariables.....	9
1.3.2. 1.4.2. Types de modèles .....	9
2. Matériel .....	11
2.1. Les Ghâts occidentaux d'Inde.....	11
2.2. Le dispositif d'Uppangala .....	11
2.2.1. Présentation du dispositif d'échantillonnage.....	11
2.2.2. Présentation de la topographie.....	12
2.2.3. Caractères principaux des espèces présentes.....	12
2.2.4. Présentation des recensements .....	12
3. Méthodes .....	15
3.1. Etape descriptive .....	15
3.2. Choix des variables susceptibles d'influencer la mortalité et de leurs transformations. ....	15
3.3. Modèle logistique et modèle linéaire généralisé.....	17
3.3.1. Fonction de lien .....	18
3.3.2. Estimation des paramètres .....	18
3.3.3. Sélection de variables.....	18
3.3.4. Ajustement du modèle .....	19
3.3.5. Validation de modèle .....	20
3.3.6. Hypothèse d'indépendance des observations .....	20
3.4. Modèle linéaire généralisé mixte.....	21
3.5. Outils .....	22
4. Résultats.....	23

4.1.	Description des variables .....	23
4.2.	Transformation des variables potentielles.....	25
4.2.1.	Accroissement .....	25
4.2.2.	Compétition.....	25
4.2.3.	Pente et exposition.....	25
4.3.	Résultats des modèles logistiques.....	26
5.	Discussion .....	29
5.1.1.	Accroissement, diamètre et compétition.....	29
5.1.2.	Pente .....	30
5.1.3.	Apport du modèle mixte .....	31
5.1.4.	Perspectives d'amélioration.....	32
	Références.....	33
	Annexes .....	37
	Annexe 1 : mortalité observée par classe de diamètre sur l'ensemble de la période de 20 ans .....	37
	Annexe 2 bis : Analyse de Ripley sur la mortalité .....	37
	Annexe 2 : comparaison de moyennes des covariables en fonction du statut Mort ou Vivant .....	38
	Annexe 3 : GLM .....	39
	Annexe 4 : Distributions des probabilités prédites par les GLM .....	40
	Annexe 5 : Courbes ROC des GLM.....	42
	Annexe 6 : GLMM .....	43
	Résumé.....	45
	Abstract .....	45

# 1. Introduction

---

Le climat et les usages de la forêt évoluent rapidement d'où la nécessité de faire aussi évoluer les pratiques sylvicoles afin d'accélérer l'adaptation des forêts. Dans ce contexte, l'élaboration de stratégies de gestion raisonnées des forêts demande d'être en mesure de faire des projections à long terme, sous différents scénarios de gestion et de forçages environnementaux, ce qui nécessite d'avoir recours à la modélisation (Clark et Clark, 1999). Dans ce contexte, il apparaît notamment important de déterminer le rôle potentiel des forêts tropicales en tant que puits ou sources de carbone (Clark, 2004) ainsi que l'impact des activités de gestion sur cet aspect.

L'influence de la diversité spécifique et de la variabilité des réponses démographiques (recrutement, croissance, mortalité) est de plus en plus mise en évidence (Clark et al., 2010). Les espèces diffèrent par leurs stratégies écologiques et sont caractérisables par une variabilité interspécifique des processus démographiques comme la croissance ou la mortalité (Westoby, 1998). Pour une espèce donnée, cette stratégie peut évoluer au cours de l'ontogénie (Clark and Clark, 1992 ; Yamada et al. 2005) et en fonction des conditions locales (Bloor and Grubb 2004). De plus, la variabilité intraspécifique de ces processus a plus récemment été identifiée comme déterminantes dans les mécanismes de coexistence des espèces (Clark, 2007, 2010 ; Vieilledent et Courbaud, 2010). La mortalité est notamment un processus clé pour expliquer les phénomènes des successions végétales (Purves et al., 2008), comprendre les mécanismes à l'origine de la distribution géographique des espèces et de la biomasse sur des échelles spatiales larges (Loehle, 1998 ; Purves, 2009 ; Delbart, 2010), mais aussi plus localement pour interpréter la structuration diamétrique d'un peuplement (Muller-Landrau et al., 2006 ; Coomes et al. 2003) et prédire la réponse d'une communauté d'arbres aux changements environnementaux et aux attaques de pathogènes (van Mantgem et Stephensen, 2007 ; Kurz et al., 2008).

## 1.1. Enjeux de la compréhension de la mortalité

La dynamique locale d'un peuplement forestier est guidée par trois grands processus : le recrutement, la croissance et la mortalité (Clark and Clark, 1999). Dans les peuplements denses, comme en forêt tropicale non perturbée, la mort d'un seul arbre de canopée ou d'un group d'arbres peut créer une trouée qui libère de l'espace laissant ainsi pénétrer la lumière dans les strates inférieures. Au cours du temps, cet espace va être colonisé, soit de manière

latérale avec le développement des houppiers d'arbres déjà en place, soit par le développement vertical de jeunes arbres déjà présents ou par une nouvelle régénération (Olivier et Larson, 1996). En fonction de la taille de cette trouée, les espèces qui remplaceront les individus morts ne seront pas forcément les mêmes, certaines espèces requièrent en effet plus de lumière pour pouvoir se régénérer, d'autres sont beaucoup plus tolérantes à l'ombre (Shugart, 1984). Le tempérament des espèces face à la quantité de lumière lors de la régénération peut résulter d'un compromis entre sa croissance et sa mortalité juvénile (Kunstler et al., 2010). Pour comprendre la dynamique de ces trouées il est donc nécessaire de pouvoir prédire la mortalité des arbres qui les génèrent et la mortalité des arbres qui s'y installent.

## **1.2. Processus biologiques et mortalité en forêt**

La mortalité des arbres en forêts est un mécanisme complexe et résulte de différents processus pouvant être indépendants les uns des autres ou, bien souvent, qui interagissent. Ces mécanismes ne sont pas systématiquement les mêmes en fonction de l'espèce observée, du stade de développement de l'individu, de son histoire de vie ou encore des conditions du milieu. Certains facteurs vont notamment affaiblir l'arbre, tandis que d'autres vont directement causer sa mort (Manion, 1981), on parle de mortalité régulière et irrégulière (Montserud, 1976). La mortalité régulière correspond à une réduction progressive de la vitalité de l'arbre qui peut résulter de la compétition pour des ressources (Peet et Christensen, 1987) ou de la sénescence. La mortalité irrégulière est plutôt décrite comme étant due à des événements aléatoires comme les attaques d'insectes ou le vent (Lee, 1971). Un affaiblissement de la vitalité de l'arbre accroît les risques de mortalité irrégulière dont les causes interagissent alors fréquemment avec des causes de mortalité régulière.

### **1.2.1. Compétition**

La mortalité des arbres les plus jeunes est en général régulière et résulte d'une compétition forte lors de leur phase d'installation (Kobe et al., 1995). L'arbre est en effet soumis à la compétition pour les ressources (lumière et espace principalement) exercée par les arbres déjà installés qui sont plus grands et possèdent un système racinaire plus développé. On parle de phénomène d'auto-éclaircie compétitive (Reineke, 1933), traduisant une mortalité densité dépendante. Ce mécanisme est relativement bien connu dans les peuplements homogènes mais semble s'appliquer dans les peuplements irréguliers et en présence de mélanges

d'espèces et notamment en forêt tropicale humide (Coomes et Allen, 2007). Reineke (1933) propose une formulation de cette loi d'auto-éclaircie liant la densité des en arbres à leurs dimension. Une des implications est d'avoir formalisée la baisse de mortalité avec l'augmentation du diamètre pour les jeunes arbres. En vieillissant, les survivants peuvent échapper à cette compétition en atteignant notamment la canopée. Différents types de compétition peuvent exister au sein d'un peuplement. On distinguera notamment la compétition symétrique de la compétition asymétrique (Pretzsch and Biber, 2010). Si deux individus différents exercent l'un sur l'autre la même compétition, la compétition est dite symétrique, si ce n'est pas le cas, la compétition est dite asymétrique. La compétition pour la lumière en forêt est par exemple plutôt de nature asymétrique puisqu'un arbre ayant atteint la canopée ne subira pas de compétition pour la lumière de la part des arbres plus petits. En revanche, la compétition racinaire est généralement considérée comme symétrique (Casper et Jackson, 1997). En forêt, bien qu'elle soit souvent ignorée, la compétition souterraine est même parfois aussi forte, voir plus forte, que la compétition pour la lumière (Casper et Jackson, 1997)

### **1.2.2. Sénescence**

La sénescence correspond à une baisse de la vitalité de l'arbre et peut être définie comme une diminution de l'efficacité dans l'utilisation des ressources liée à une diminution de l'efficacité respiratoire ou de conductance hydrique (Gower et al., 1996 ; Hubbard et al., 1999). La mortalité est alors parfois décrite par certains auteurs (Vieilledent et al., 2009 ; Chao et al., 2008 ; Coomes et al., 2007) comme ayant une forme en U : la mortalité est forte pour les faibles diamètres (mécanisme en jeu = compétition) puis décroît et ré-augmente pour les gros diamètres (mécanisme en jeu = sénescence). La mortalité exprimée comme une fonction de la taille peut donc refléter la compétition juvénile et la sénescence (Harcombe, 1987).

### **1.2.3. Chute d'arbre : le chablis**

La mortalité peut être liée à différents processus biologiques mais des causes physiques de mortalité sont aussi observées, notamment le chablis (Durrieu de Madron, 1993.). Un chablis correspond à la chute de la partie aérienne d'un arbre, pouvant aussi entraîner une partie de son appareil racinaire.

Tous les arbres ne présentent pas le même risque de chablis, en effet un arbre de grande dimension et dont le houppier est déséquilibré (forte dissymétrie par exemple) est plus

sensible au vent. De même un arbre dont le système racinaire présente des faiblesses (caractéristique de l'espèce ou condition locales du sol) est moins stable (Canham, 2001). Il est alors reconnu que les fortes pentes augmentent le risque de chablis en favorisant une croissance dissymétrique des houppiers, en offrant parfois des conditions moins bonnes de sol et en exposant plus fortement les arbres au vent (Moravie, 1999 ; Durieu de Madron, 1993). Les chablis touchent de manière directe en majorité les arbres de grande dimension, on parle de chablis primaire. Cependant la chute d'un gros arbre entraîne en général les arbres situés en dessous, on parle de chablis secondaire.

## **1.3. Modélisation de la mortalité en forêt**

### **1.3.1. 1.4.1. Prise en compte des différentes covariables**

La croissance est une mesure intégrative de la vitalité des arbres (Bigler and Bugmann 2003). Elle constitue un indicateur synthétique de plusieurs causes régulières de mortalité (dont la compétition et l'âge notamment) et est fréquemment utilisée en modélisation (Vieilledent et al., 2010). Dans le cas où les arbres sont spatialisés, il est fréquent d'utiliser un indice de compétition dépendant du voisinage de l'arbre (Olono et al., 2009).

### **1.3.2. 1.4.2. Types de modèles**

Les modèles de mortalité sont souvent mis en œuvre pour intégrer des modèles plus généraux de la dynamique forestière (Hawkes, 2000). Il en existe deux grandes catégories : les modèles mécanistes (basés sur les processus physiologiques) et, beaucoup plus fréquemment utilisés, les modèles empiriques (obtenus en ajustant des relations entre prédicteurs observables sur le terrain et la mortalité) (Hawkes, 2000). L'étude présentée dans ce mémoire est basée sur une approche empirique de la modélisation de la mortalité à partir de données récoltées sur le dispositif permanent d'Uppangala située dans une forêt tropicale du sud ouest de l'Inde.

Différentes approches sont classiquement utilisées pour modéliser empiriquement la mortalité. L'analyse de survie diffère notamment des autres approches en permettant de modéliser le temps d'occurrence d'un événement (Collett, D., 2003). Elle a été développée et beaucoup utilisée en médecine notamment dans le cadre de prédiction du temps de survie de patients en fonction de différents traitements ou variables continues caractérisant les individus (Collet, 2003). Ces analyses prennent en compte explicitement la répétition dans le temps des observations sur un même individu et le fait que les observations soient incomplètes (on parle

de données « censurées »). Dans le cas forestier, l'évènement de mortalité n'est observé que pour une très faible part des individus suivis. Les arbres ont, en effet, une durée de vie bien supérieure à la durée des observations, ce qui en fait un sujet d'étude délicat.

Des approches alternatives modélisent directement la probabilité de mort sur un pas de temps. La régression logistique est notamment beaucoup utilisée. Ce type de régression désigne une forme de la relation entre des variables explicatives et la variable à expliquer (forme de type sigmoïde). Cette relation logistique entre variables expliquées et explicatives a été proposée par Verhulst en 1838 comme résultat de la résolution d'un système d'équations de dynamique d'une population. Cette solution s'est avérée bien décrire l'évolution de la taille d'une population et a largement été reprise en écologie (Schtickzelle, 1981). Il a par la suite été montré sur des modèles théoriques que la mortalité pouvait dans certains cas être une fonction logistique de covariables. Ainsi introduite dans le contexte forestier par Hamilton (1974), la régression logistique a été depuis largement utilisée pour y modéliser la mortalité (Hamilton et Edwards, 1976 ; Hamilton, 1986 ; Yao et al., 2001 ; Yang et al., 2003 ; Zens et Peart, 2003 ; Chao, 2008 ; Lines et al. 2010 ; Rüger et al., 2011).

Dans cette étude, plusieurs modèles logistiques seront développés pour étudier la mortalité. Le modèle linéaire généralisé (GLM) est une technique d'ajustement du modèle logistique (McCullagh and Nelder, 1989) et a été utilisé dans cette étude. L'utilisation d'un modèle mixte a été nécessaire dans cette étude pour prendre en compte les mesures répétées sur les arbres du dispositif (Lee et al., 2006 ; Zuur et al., 2009).

## 2. Matériel

---

### 2.1. Les Ghâts occidentaux d'Inde

Le dispositif forestier permanent d'Uppangala, mis en place et suivi par l'Institut Français de Pondichéry, est localisé dans une forêt non perturbée des Ghâts occidentaux d'Inde<sup>1</sup>. Il se trouve sur la partie haute des contreforts des Ghâts (signifie « escaliers » ou « marches » en langue Hindi) qui marque la limite avec le plateau du Karnataka. Cette zone de l'Inde est notamment caractérisée par un relief contrasté et un climat tropical humide (Pélissier et al. 2011).

Le dispositif d'Uppangala est situé en forêt tropicale humide dans le « Pushpagiri Wildlife Sanctuary » qui constitue une des reliques de forêt naturelle en Inde. Le site est strictement protégé depuis 1990 et aucune perturbation humaine majeure n'y est connue. Les seules perturbations connues sont limitées à quelques prélèvements de produits forestiers réalisés par des habitants de villages proches (Salaün, 1995). Uppangala, le plus proche village, est situé à 10 km du dispositif et tout accès motorisé était impossible avant l'ouverture d'une piste dans les années 1980 (Pélissier et al., 2011).

Cette zone géographique est caractérisée par une forte saisonnalité avec un pic de pluviométrie durant la période de mousson et une saison sèche hivernale marquée (Pascal and Pélissier 1996). 90% des 5000 mm de pluies annuelles sont reçus durant la saison humide (de Juin à Octobre). La saisonnalité marquée influence d'ailleurs la croissance de certaines espèces sur le dispositif (Madelaine-Antin, 2009). Peu de dispositifs d'étude permanents dédiés à l'étude des forêts tropicales humides sont aussi marqués qu'Uppangala par la saisonnalité (Elouard et al. 1997).

### 2.2. Le dispositif d'Uppangala

#### 2.2.1. Présentation du dispositif d'échantillonnage

Le dispositif est composé de 5 bandes orientées Nord – Sud de 20 mètres de large et d'une longueur comprise entre 180 et 370 mètres (parcelles A, B, C, D et E). Elles sont espacées de 100m. Le dispositif est aussi composé de 5 parcelles rectangulaires (parcelles H1, H2, H3, R et S) qui recourent les 5 bandes. L'ensemble des parcelles représente une superficie de 5,17ha. Tous les arbres sont identifiés et géolocalisés.

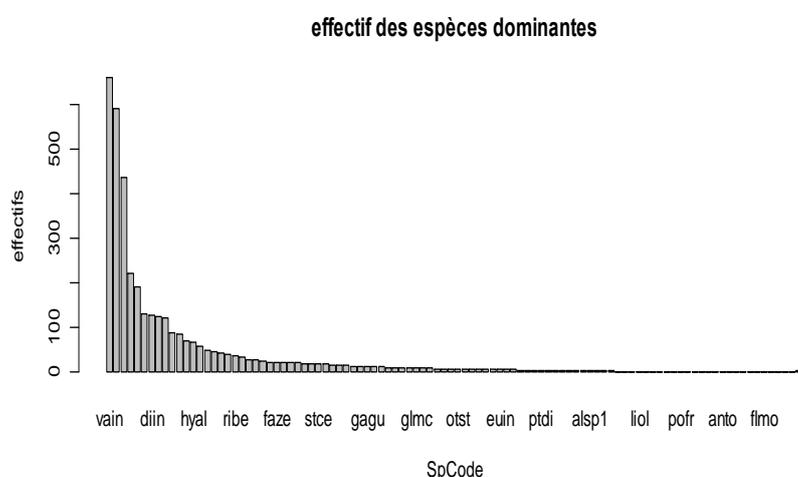
---

<sup>1</sup> Un des 42 points chauds de biodiversité mondiaux (<http://www.biodiversityhotspots.org/Pages/default.aspx>)

### 2.2.2. Présentation de la topographie

L'altitude du dispositif est comprise entre environ 400 et 600 mètres. La pente principale est orientée nord-sud et des ruisseaux temporaires alimentés en saisons des pluies ont creusé des vallons créant globalement une opposition de versant est-ouest. Des mesures de pentes ont été réalisées sur les arrêtes des quadrats de 10m X 10m sur tout le dispositif (détails dans Pélissier et al., 2011) et donnent ainsi accès à un attribut de pente et d'exposition pour chaque arbre.

### 2.2.3. Caractères principaux des espèces présentes



Cent deux espèces d'arbres sont présentes sur le dispositif, mais le peuplement est largement dominé en termes d'effectifs et de surface terrière par quelques espèces constituant l'ossature de la forêt et ayant un rôle bien individualisé dans la stratification des espèces (Pélissier 1995, Elouard et al.

1997) : *Dipterocarpus indicus*, *Vateria indica*, *Myristica dactyloïdes*, *Knema attenuata* et

Figure 1 : abondance des espèces présentes sur le dispositif d'Uppangala

*Humboldtia brunonis*. Ces espèces représentent plus de 50 % des effectifs et de la surface terrière du peuplement. Des travaux de Pascal (1988) mis à jour par Ramesh et al. (2010) ont mis en évidence des grands ensembles structuraux et classé les espèces présentes sur le dispositif en quatre groupes (strates potentiellement atteignables par les individus de l'espèce) : une strate émergente, une strate canopée, une strate sous-canopée et une strate sous bois. Ces quatre groupes comportent des effectifs proches sur le dispositif (respectivement : 903, 826, 919 et 1189 individus). Ces groupes contiennent en revanche un nombre relativement variable d'espèces (respectivement : 7, 32, 24 et 38 espèces). Chacun de ces ensembles structuraux est représenté par au moins une des 5 espèces majoritaires.

### 2.2.4. Présentation des recensements

parcelle	surface (ha)	date d'installation	parcelle	surface (ha)	date d'installation
A	0.36	0	H3*	0.12	1341
B	0.74	0	H=(H1+H2+H3)*	1.77	170
C	0.66	0	R	0.24	1341
D	0.68	0	R*	0.72	1030
E	0.68	0	S	0.12	1341
H1*	1.1	0	S*	0.24	720
H2*	0.55	762			

Tableau 1 : Surface et dates d'installation des parcelles échantillonnées sur le dispositif d'Uppangala. Les dates sont en jours depuis le 1<sup>er</sup> Mars 1990 (correspondant à la date des premières mesures de diamètres). \* correspond à la partie de la parcelle qui chevauche les parcelles A, B, C, D et E (d'après Pélissier et al., 2011).

L'installation des 5,17 ha du dispositif s'est déroulée sur une période de 4 ans. L'installation des parcelles A, B, C, D et E entre 1989 et 1990 a été suivie par la mise en place des parcelles H, R et S entre 1990 et 1993 (tableau 1).

### 2.2.4.1. Mesure des diamètres

Le diamètre du tronc à 1,30 m (diameter at breast height = dbh) est mesuré à l'aide de dendromètres en ruban (figure 2) fixés de manière permanente sur les troncs. La précision théorique des dendromètres est 0.02 mm (Pélissier et al., 2011) et la mesure est toujours réalisées exactement au même endroit sur le tronc, ce qui réduit les incertitudes de mesure par rapport à des dispositifs de mesure qui ne seraient pas fixés de manière permanente.



Figure 2: Dendromètre en ruban

A partir de l'installation du dispositif et durant les 6 premières années (1990 à 1995), deux recensements ont été réalisés par an (un avant et un après la période de mousson). Un recensement par an a par la suite été effectué jusqu'à aujourd'hui (tableau 2). Ainsi, 27 campagnes de mesures ont donc été conduites depuis l'installation du dispositif.

dates	dates (jours)	dates	dates (jours)	dates	dates (jours)
15/03/1990	0	15/11/1994	1705	15/04/2003	4779
15/11/1990	244	15/04/1995	1857	15/04/2004	5144
15/04/1991	396	15/11/1995	2070	15/04/2005	5509
15/11/1991	609	15/11/1996	2435	15/04/2006	5874
15/04/1992	761	15/11/1997	2800	15/04/2007	6240
15/11/1992	974	15/11/1998	3166	15/04/2008	6605
15/04/1993	1126	15/10/1999	3500	15/04/2009	6970
15/11/1993	1339	15/12/2000	3926	15/04/2010	7335
15/04/1994	1491	15/12/2001	4292	15/04/2011	7700 <sup>2</sup>

Tableau 2 : Dates des campagnes de mesure. En vert : mesure après la mousson. En bleu : mesure avant la mousson.

<sup>2</sup> Les données du dernier inventaire doivent encore être ajoutées à la base de données et ne sont donc pas encore disponibles.

L'utilisation de dendromètres fixes permet une très grande précision des mesures. Cependant, la variation des dates de mesure entraîne une variabilité des accroissements mesurés. La mesure biannuelle de la croissance permet de mettre en évidence un nombre élevé d'accroissements négatifs pendant la saison sèche (novembre-avril), dus à la contraction du tronc (Moravie, 1996). Il est réaliste de penser qu'un effet sur la mortalité peut être observé et pour ne pas introduire cet effet intra-annuel dans les modèles, les analyses porteront donc sur un seul relevé annuel (mesure d'après mousson jusqu'en 2001 et mesures d'après mousson à partir de 2003).

#### ***2.2.4.2. Mesures de la mortalité***

Il arrive que le diamètre ne soit pas mesurable (arbre cassé ou tombé par exemple). Pour des raisons d'homogénéité des données récoltées, quand l'arbre est déclaré mort, son diamètre n'a pas été enregistré. Aucune mesure de diamètre n'est donc disponible quand l'arbre est reconnu mort.

#### ***2.2.4.3. Mesures du recrutement***

Tous les 4 ou 5 ans, les arbres recrutés (ayant franchi le seuil de 10 cm diamètre) sont équipés d'un dendromètre, identifiés et localisés. Ces arbres sont ne sont donc suivis qu'à partir de leur date de recrutement et intégrés dans la base de données avec un code particulier pour la première année de mesure. Les campagnes de recrutement ont eu lieu en 1994, 1998, 2002, 2007 et 2010.

## 3. Méthodes

Tout au cours de cette étude, les analyses ont été réalisées à différents niveaux de description : au niveau du peuplement tout entier, au niveau de groupes d'espèces identifiées comme structurantes des strates (émergente, canopée, sous-canopée et sous-bois représentées respectivement par les codes 0, 1, 2, 3), ainsi qu'au niveau des 5 espèces identifiées comme fortement présentes et/ou structurantes au sein de ces strates (tableau 3).

	Effectif total	Nombre d'arbres mort sur la période	Strate potentielle
<i>Vateria indica</i>	659	62	0 = émergente
<i>Humboldtia brunonis</i>	590	82	3 = sous-bois
<i>Myristica dactyloïdes</i>	435	60	2 = sous-canopée
<i>Knema attenuata</i>	222	26	2 = sous-canopée
<i>Palaquium ellipticum</i>	192	31	1 = canopée

Tableau 3 : effectifs des espèces représentant chaque strate potentielle

### 3.1. Etape descriptive

L'influence des variables potentiellement explicatives sur les données font dans un premier temps l'objet d'analyses descriptives simples (comparaison de moyennes, distribution diamétriques des variables en fonction du statut mort / vivant). L'effet de la date a été évalué grâce à l'estimation de taux de mortalité annuelle (différence de diamètre entre les deux dates / temps écoulé entre les deux dates). Enfin, une analyse de Ripley a été conduite pour dégager une éventuelle structuration spatiale des arbres morts et des arbres vivants en fonction de la pente. Cette analyse permet d'évaluer la densité de voisins dans un rayon croissant autour de chaque arbre (Ripley, 1977) et d'évaluer les structures de répulsions ou d'agréations spatiales.

### 3.2. Choix des variables susceptibles d'influencer la mortalité et de leurs transformations.

Différents auteurs montrent que les variables prises en compte dans les modèles de mortalité en forêts sont parfois plus explicatives après transformation (Yang et al. 2003 ; Rüger et al., 2009). On ne connaît en effet pas la forme exacte de la relation entre ces covariables et la mortalité, mais pour tester plusieurs hypothèses et pour rester dans un cadre paramétriques, plusieurs transformations de variable ont été testées et sont ici présentées.

Dans le tableau 4, différents indices sont proposés. Les indices 1 et 2 sont symétriques tandis que les indices 3, 4, 5, 6 et 7 sont asymétriques. La surface terrière a soit été ramenée au diamètre de l'arbre central soit au carré de ce diamètre. Le fait de ramener la surface terrière au diamètre de l'arbre a pour objectif de rendre l'indice asymétrique. Cet indice dépend ainsi non seulement de l'entourage de l'arbre, mais aussi de ses propres dimensions. La surface terrière a été soit ramenée au diamètre ou au diamètre au carré pour tester l'importance relative du voisinage et des caractéristiques de l'arbre.

n°	Nom	Descriptif
1	N5, N10, N15	Effectif dans un rayon de 5, 10 ou 15m autour de l'arbre.
2	g5, g10, g15	Surface terrière (m <sup>2</sup> /ha) dans un rayon de 5, 10 ou 15m autour de l'arbre.
3	G5, G10, G15	Surface terrière (m <sup>2</sup> /ha) des arbres plus gros dans un rayon de 5, 10 ou 15m autour de l'arbre
4	g5surDBH, g10surDBH, g15surDBH	Surface terrière (m <sup>2</sup> /ha) dans un rayon de 5, 10 ou 15m autour de l'arbre divisée par le diamètre de l'arbre.
5	g5surDBH2, g10surDBH2, g15surDBH2	Surface terrière (m <sup>2</sup> /ha) dans un rayon de 5, 10 ou 15m autour de l'arbre divisée par le diamètre de l'arbre au carré.
6	G5surDBH, G10surDBH, G15surDBH	Surface terrière (m <sup>2</sup> /ha) des arbres plus gros dans un rayon de 5, 10 ou 15m autour de l'arbre divisée par le diamètre de l'arbre.
7	G5surDBH2, G10surDBH2, G15surDBH2	Surface terrière (m <sup>2</sup> /ha) des arbres plus gros dans un rayon de 5, 10 ou 15m autour de l'arbre divisée par le diamètre de l'arbre au carré.

**Tableau 4 : Indices de compétition testés comme variables explicatives de la mortalité (un rayon de compétition classiquement proposé en forêt tropicale en 10 (Nath et al., 2011), 5, 10 et 15 ont été ici testés)**

De même, différentes transformations de la variable accroissement ont été testées. La transformation log est utilisée par certains auteurs (e.g. : Rüger et al., 2011) mais suppose que les valeurs d'accroissement sont toutes positives, ce qui n'est pas le cas ici. Cette transformation exacerbe les différences d'accroissement pour les faibles valeurs de cette variable alors que les différences sont amoindries pour les fortes valeurs. Dans cet esprit, d'autres transformations permettent de prendre en compte les accroissements négatifs : la transformation racine cubique ou de type  $x \sim x^{1/k}$  (avec k impaire). Les valeurs 3, 5 et 7 ont été testées. La pente a soit été considérée comme variable quantitative (pente faible vs pente forte avec un seuil de 50% de pente).

Les covariables potentielles sont testées une à une et comme seule covariable pour expliquer la mortalité. Le choix des variables les plus susceptibles d'avoir un effet sur la mortalité est réalisé à l'aide de critères d'information pénalisés comme l'AIC (Akaike's Information Criterion) ou le BIC (Bayesian Information Criterion), qui sont basés sur la vraisemblance des

observations par rapport au modèle et pénalisés par le nombre de paramètres (équation 1). L'objectif recherché par l'utilisation de ces critères est de choisir un modèle parcimonieux (compromis entre qualité d'ajustement et nombre de paramètre).

$$AIC = 2 \times k - 2 \times \log(L) \qquad BIC = k \times \log(n) - 2 \times \log(L)$$

*n = nombre d'observations*  
*k = nombre de paramètres estimés par le modèle*  
*L = vraisemblance du modèle*

*Equation 1 : formules des différents critères utilisés*

### 3.3. Modèle logistique et modèle linéaire généralisé

Dans le cadre des modèles paramétriques, le choix de la technique de modélisation repose en partie sur les hypothèses faites sur la forme de la relation existante entre la variable expliquée et les variables explicatives. Un modèle logistique traduit un lien non linéaire entre des variables explicatives et une variable expliquée (de forme sigmoïde pas exemple). Dans le cas d'une variable binaire (c'est-à-dire l'espérance de la variable), on cherche à modéliser la probabilité d'occurrence d'un des deux événements, et c'est cette probabilité que l'on cherche à modéliser par une fonction logistique de covariables explicatives. Le modèle linéaire généralisé est alors utilisé dans ces situations comme technique d'ajustement du modèle logistique.

$$g(E[Y]) = \sum_i \alpha_i X_i$$

Y est la variable à expliquer (variable binaire prenant les valeurs 0 (survie) ou 1 (mort))

E[Y] correspond donc à la probabilité de mort : on la notera p

X<sub>i</sub> est la i<sup>ème</sup> covariable du modèle

α<sub>i</sub> est le i<sup>ème</sup> coefficient du modèle

$$g(p) = \text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \sum_i \alpha_i X_i \Rightarrow p = \frac{e^{\sum_i \alpha_i X_i}}{1 + e^{\sum_i \alpha_i X_i}} \in ]0, 1[$$

*Equation 2 : formalisation du modèle linéaire généralisé dans le cas d'une variable binomiale*

Le choix d'une fonction de lien est la première étape indispensable à la mise en œuvre de cette technique.

### 3.3.1. Fonction de lien

Cette fonction sert à linéariser la relation entre réponse et prédicteur pour se rapprocher du cadre de la régression linéaire classique. On cherche à estimer une probabilité qui est donc comprise entre 0 et 1. Le plus simple est de trouver une fonction de lien bijective de  $]0 ; 1[$  vers  $]-\infty ; +\infty[$  qui permet d'e s'assurer que les probabilités prédites seront comprises entre 0 et 1. Dans le cas d'une régression logistique où la variable suit une loi binomiale, on utilise en général la fonction de lien  $\text{logit}(p) = \log(p/1-p)$  (équation 2).

### 3.3.2. Estimation des paramètres

Du fait de la généralisation, on n'estime plus les paramètres du modèle selon le critère des moindres carrés mais par estimation au maximum de vraisemblance. En effet, l'utilisation des moindres carrés suppose de faire l'hypothèse de normalité et d'homoscédasticité des résidus trop forte. Dans le cas d'une variable binomiale, on utilise le critère du maximum de vraisemblance (équation 3) que l'on estime par un algorithme itératif de type Expectancy Maximisation (McCullagh and Nelder, 1989) appliquée à la maximisation de la vraisemblance restreinte.

$$\log(V) = \sum_j [\log(p_j) - \log(1 - p_j)] = \sum_j \text{logit}(p_j) \quad \text{avec} \quad p_j = \frac{e^{\sum_i \alpha_i X_{ij}}}{1 + e^{\sum_i \alpha_i X_{ij}}}$$

V est la vraisemblance des observations sur tous les individus j suivant les paramètres  $\alpha_i$  estimés :

On pourra remarquer que la fonction de lien apparaît logit apparaît dans l'expression de la vraisemblance : il s'agit de la fonction de lien «canonique»

*Equation 3 : formule de la log vraisemblance*

### 3.3.3. Sélection de variables

Dans le cas du modèle généralisé, il a été montré que la statistique de Wald (équation 4) suivait une loi asymptotique du Chi2 par application du théorème central limite (Diggle et al., 1994), ce qui permet tester l'hypothèse de nullité du paramètre.

$$W = \frac{\bar{\alpha}_i^2}{\bar{V}(\alpha_i)} \sim \chi^2(1) \quad \text{asymptotiquement sous l'hypothèse } H_0 : \bar{\alpha}_i = 0$$

*Equation 4 : test de Wald*

Une alternative à l'utilisation de la statistique de Wald, qui dépend beaucoup de la validité de l'hypothèse faite sur la loi de l'estimateur du paramètre testé, est l'utilisation du rapport de vraisemblances du modèle avec le paramètre et du modèle sans le paramètre. Il semble en effet plus fiable d'après certains auteurs comme Agresti (1996). Le logarithme de ce rapport de vraisemblances suit aussi une loi du Chi<sup>2</sup> et permet ainsi de tester si la vraisemblance du modèle avec le paramètre est significativement meilleure que celle du modèle sans.

Des procédures de sélection automatique de variable sur la base du critère AIC (équation 1) permettent une première approche. Certains auteurs remarquent que l'utilisation de l'AIC comme critère peut conduire à favoriser les modèles complexes (Link & Barker 2006). Dans certains cas, la sélection de variables uniquement sur la base de l'AIC sélectionne des variables non significatives. Une sélection manuelle vient alors compléter cette approche. Si l'écart d'AIC est très faible et que l'approche descriptive n'a pas mis en évidence l'intervention de la variable en question dans le phénomène, la variable ne sera pas conservée. En revanche s'il semble à priori que la variable a de l'importance mais qu'elle n'a pas été sélectionnée par l'AIC, elle sera finalement rajoutée au modèle et son influence testée par un test du rapport de vraisemblance. En effet, un modèle reflétant les mécanismes biologiques connus sera préféré à un modèle sélectionné uniquement pour son ajustement aux données. Il sera plus facilement interprétable sur le jeu de données utilisé pour l'estimation des paramètres et permettra un transfert plus simple vers d'autres jeux de données (Hamilton, 1986).

#### **3.3.4. Ajustement du modèle**

Dans le cas d'une variable expliquée continue, l'ajustement d'un modèle peut être évalué par une mesure de type corrélation entre réalisation de la variable aléatoire étudiée et estimation de son espérance prédite par le modèle. Néanmoins ce schéma ne s'applique pas dans le cas d'une variable expliquée binaire. Il est en revanche possible d'utiliser le test de Hosmer et Lemeshow (1988, 1989) basé sur la statistique du  $\chi^2$  dans ce cadre de régression linéaire généralisée. Le principe est de catégoriser les variables explicatives (indépendantes) continues et de calculer une statistique du  $\chi^2$  pour chaque classe. Il faut ensuite comparer la valeur calculée à la valeur critique d'une table du  $\chi^2$  afin de vérifier si la prédiction du modèle n'est pas significativement différente de ce qui est observé dans les données (Yang et al., 2003).

Si le nombre de covariables est important, les classes créées ne comporteront que peu d'observations et même si l'ajustement est bon, les statistiques du  $\chi^2$  calculées inciteront à rejeter le modèle (Hosmer et al., 1997). En effet, dans le cas du modèle linéaire généralisé, on prédit la probabilité de mortalité qui est prédite et pas de l'évènement lui-même. Si on code l'évènement binaire par 1 ou 0, c'est l'espérance de cet évènement qui est prédite. Pour approcher cette espérance, il faut un nombre de répétitions suffisamment élevé et donc des effectifs suffisants dans les classes créées quand l'évènement est rare comme la mortalité. Ce test est trop dépendant de la manière dont les classes sont formées et n'est pas puissant (Agresti, 2002).

### **3.3.5. Validation de modèle**

La significativité globale du modèle est évaluée en utilisant le rapport de vraisemblance (Hosmer et Lemeshow, 2000). De plus l'analyse de ROC (Receiver Operating Characteristic) permet de tester la régression logistique en évaluant la sensibilité (taux de vrais positifs) et la spécificité du modèle (taux de faux positifs). Elles sont souvent utilisées en statistiques pour montrer les progrès réalisés grâce à un classifieur ou prédicteur binaire (prédiction du statut mort ou vivant ici) lorsque le seuil de discrimination varie (prédiction déterministe du statut à partir d'un seuil sur la probabilité de mort). Cette analyse permettra ici d'évaluer la qualité de prédiction du modèle sur les données qui ont servi à le calibrer. Il s'agit donc d'une sorte de mesure d'ajustement du modèle.

### **3.3.6. Hypothèse d'indépendance des observations**

Comme de nombreuses méthodes de modélisation le GLM demande de faire une hypothèse d'indépendance des observations pour pouvoir interpréter le modèle. Cette hypothèse est raisonnable à considérer si on ne prend en compte qu'une observation par arbre. Dans un premier temps la mortalité a donc été modélisée à partir de la dernière observation disponible pour chaque arbre. Mais afin d'exploiter au mieux les données disponibles, et prendre en compte toutes les observations, il faut aussi prendre en compte le fait qu'un même individu est observé plusieurs fois et que ces observations ne sont donc pas indépendantes. L'utilisation d'un modèle mixte avec effet individuel est une méthode qui s'est particulièrement développées depuis quelques années pour traiter ce type de données à mesures répétées (Bolker et al., 2009 ; Lee et al., 2006 ; Zuur et al., 2009) et a été mis en œuvre ici.

### 3.4. Modèle linéaire généralisé mixte

Un modèle mixte est un modèle qui prend en compte à la fois des effets fixes et des effets aléatoires. Ce concept qui a été développé sur les modèles linéaires classiques a plus récemment été adapté au modèle linéaire généralisé (Hadfield, 2010). On appelle modèle linéaire généralisé mixte (GLMM) un modèle linéaire généralisé dans lequel le prédicteur linéaire est une combinaison linéaire de paramètres (pour les effets fixes) et de variables aléatoires (pour les effets aléatoires) (équation 5).

$$g(E[Y]) = \sum_i \alpha_i X_i + \sum_j \mu_j$$

Y est la variable à expliquer (variable binaire prenant les valeurs 0 (survie) ou 1 (mort))

E[Y] correspond donc à la probabilité de mort : on la notera p

$X_i$  est la  $i^{\text{ème}}$  covariable du modèle

$\alpha_i$  est le  $i^{\text{ème}}$  coefficient du modèle

$\mu_j$  est le  $j^{\text{ème}}$  effet aléatoire  $\mu_j \sim$

$$g(p) = \text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \sum_i \alpha_i X_i \Rightarrow p = \frac{e^{\sum_i \alpha_i X_i}}{1 + e^{\sum_i \alpha_i X_i}} \in ]0, 1[$$

#### Equation 5 : formalisation du modèle linéaire généralisé dans le cas d'une variable binomiale

Une approche bayésienne a ici été choisie pour l'estimation des paramètres. Elle permet de s'affranchir des hypothèses de normalité. Les paramètres estimés pour les effets fixes sont comparés aux quantiles de leur distribution postérieure ce qui permet d'évaluer leur significativité.

Dans le modèle mixte, il est possible de spécifier la structure de la matrice de variance covariance pour chaque effet aléatoire. On pourra par exemple choisir une matrice diagonale indiquant que les covariances sont nulles. L'hypothèse ici faite est que l'observation faite sur un arbre ne dépendant pas de celles faites sur ses voisins mais que des observations faites sur un même arbre ne sont pas indépendantes. Au contraire on peut spécifier que les prédictions faites sur un individu dépendent de celles faites sur ses voisins (en cas de constatation de structure spatiale du processus modélisé par exemple). Dans ce cas, différentes structures de matrice de variance covariance sont possibles en fonction du type de structure observée (agrégation ou répulsion spatiale par exemple).

Les variables sélectionnées par l'approche GLM sont reprises pour y ajouter les effets aléatoires. Pour vérifier que ces variables sont bien les plus influentes mêmes quand les effets

aléatoires sont pris en compte, les variables restantes ont été ajoutées à tour de rôle et les modèles comparés en terme de d'AIC et de BIC.

Il n'existe pas de test statistique général pour évaluer la significativité de l'effet aléatoire pris en compte, en effet, l'hypothèse nulle correspondant à ce type de tests serait que la variance de la variable décrivant l'effet aléatoire est nulle. Une variance étant par définition positive, cette valeur nulle correspondrait donc à une borne de l'ensemble des valeurs qu'elle peut prendre. Dans ce cas, aucune statistique de test suivant une loi connue n'est associée à ce paramètre sous l'hypothèse nulle (Lee et al., 2006). Il est en revanche possible d'évaluer la qualité générale du modèle en se basant sur les analyses proposées pour le modèle à effets fixes uniquement. On n'utilise ici le jeu de données complet où les observations sont répétées sur les individus. Les AIC et BIC de ces modèles ne sont alors pas comparables à ceux obtenus par les GLM précédant car le nombre d'observations est bien plus grand. Pour évaluer le gain apporté par l'effet aléatoire, on réalise une comparaison des critères AIC et BIC entre modèle avec effet aléatoire et sans effet aléatoire où le modèle sans effet aléatoire est ajusté sur le jeu de données complet. Les effets arbres et les effets date ont été ici testés.

### 3.5. Outils

Toutes les analyses statistiques ont été réalisées à l'aide du logiciel libre R et grâce aux différents packages disponibles (R Development Core Team, 2010).

# 4. Résultats

## 4.1. Description des variables

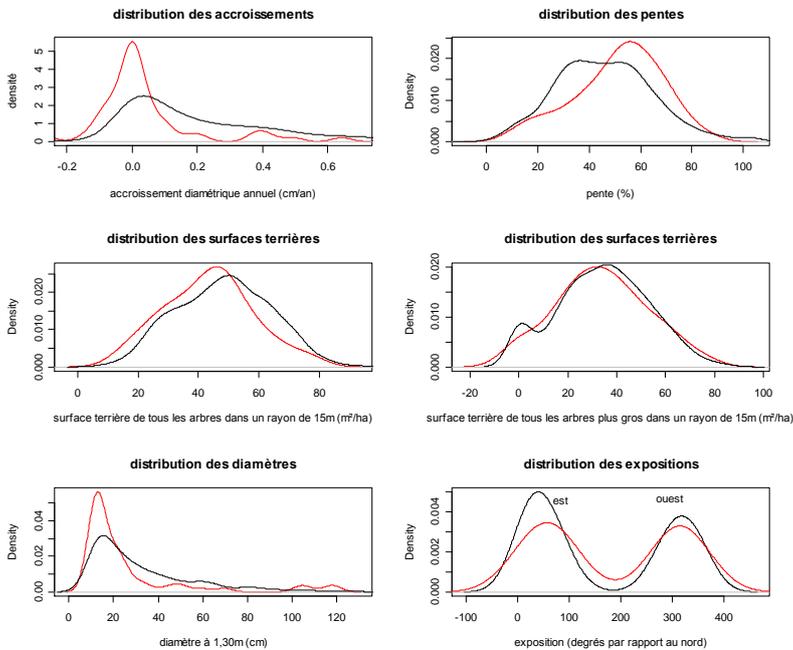


Figure 3 : analyse des criptives des variables disponible pour *Vateria indica* (les graphiques pour les autres espèces les différentes strates et le peuplement sont disponibles en annexe)

Au niveau du peuplement entier l'analyse graphique (graphique similaire à la figure 3) met en évidence des un accroissements moyens plus faibles, une pente plus forte et une surface terrière (g15) plus faible pour les arbres morts. Ce qui est confirmé par les comparaisons de moyennes (p-value : 0,00008 ; <0,000001 ; 0,00054). g10 et g15 sont les deux seuls indices de compétition dont la moyenne est différente (p-value : 0,04532 ; 0,00054) toutes strates confondues. Aucune différence sur les diamètres moyens

ou les expositions n'est constatée.

Pour toutes les strates et toutes les espèces analysées, on observe un accroissement moyen significativement plus faible pour les arbres morts, la pente est en moyenne significativement plus forte pour les arbres morts dans le cas de la strate émergente, la strate de canopée, de la strate de sous canopée et de *Vateria indica*. Le diamètre des arbres morts est en moyenne plus élevé pour *Palaquium ellipticum* et pour la strate de sous bois ( 28,6 contre 18,1cm avec une p-value de 0,01338 et 15,5 cm contre 14,3 cm avec une p-value de 0,01862). On constate par ailleurs en observant les distributions diamétriques de la mortalité (annexe 1) que pour *Vateria indica*, la mortalité diminue d'abord avec le diamètre puis augmente pour les gros diamètres et que pour *Humboldtia brumonis* la mortalité augmente avec le diamètre. Une différence d'exposition moyenne entre les arbres morts et les arbres vivants est observée pour seulement pour la strate 1.

Pour la strate de canopée, *Knema attenuata* et *Humboldtia brumonis*, la compétition moyenne au niveau des arbres morts et des arbres vivants n'est pas statistiquement différente (test de Student en Annexe 2).

Le tableau 5 synthétise le type d'indice dont la valeur moyenne est différente entre les arbres morts et vivants. Dans le cas de la strate émergente et de *Vateria indica* qui en fait partie, la compétition observée est en moyenne plus forte au niveau des arbres morts que des arbres vivants. En revanche pour les strates inférieures et pour *Palaquium ellipticum* et *Myristica dactyloïdes*.

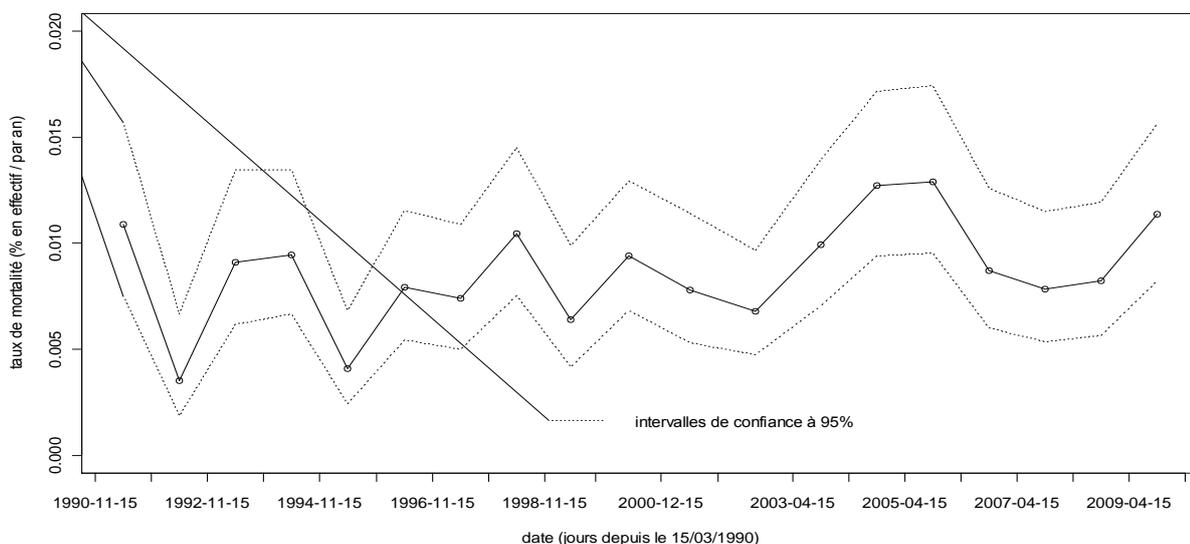
	Strate émergente	Strate sous canopée	Strate sous bois	Vateria indica	Myristica dactyloïdes	Palaquium ellipticum
Type d'indice	/DBH	15m	tous	/DBH	15m	10 m et 15 m
Compétition la plus forte	morts	vivants	vivants	Morts	vivants	vivants

**Tableau 5 : influence de la compétition. Le type d'indice reflète soit le rayon pour lequel la compétition entraîne des différences significatives (voir annexe 2) de moyenne en fonction du statut morts/vivants. La ligne "compétition la plus forte" indique si la compétition est plus forte pour les arbres morts ou les arbres vivants.**

On constate sur la figure 4 que le taux de mortalité annuel calculé entre chaque recensement est variable. On pourra prendre en compte un effet date dans les modèles qui suivent. Enfin, une analyse de Ripley sur les arbres morts a permis de mettre en évidence l'existence d'une structuration spatiale agrégative de la mortalité (annexe 2 bis : la fonction L de Ripley calculée à partir des données est au dessus de l'intervalle de confiance obtenu par simulation dans des semis de points aléatoires).

### évolution de la mortalité annuelle sur 20 ans (toutes espèces confondues)

calculs de mortalité réalisés sur une année (entre novembre et novembre ou entre avril et avril)



**Figure 4 : mortalité dans le temps sur tout le peuplement. Les dates correspondent aux recensements et la mortalité est calculée entre deux recensements. Les intervalles de confiance sont calculés en utilisant l'approximation gaussienne de la loi Binomiale.**

## 4.2. Transformation des variables potentielles

### 4.2.1. Accroissement

Parmi l'accroissement relatif (expgr) et l'accroissement absolu (incgr), c'est l'accroissement absolu qui a été retenu par les critères AIC et BIC à l'exception de *Palaquium ellipticum* et *Myristica dactyloïdes* pour lesquelles l'accroissement relatif l'a emporté. Néanmoins, les différences en termes d'AIC ou BIC sont faibles comparé à la réduction d'AIC par rapport au modèle nul, et pour rendre les modèles plus comparables, c'est l'accroissement absolu qui a été retenu. Parmi les différentes transformations de l'accroissement qui ont été de la même manière testées, c'est la racine cubique de l'accroissement (incgr0.33) qui a systématiquement le plus réduit l'AIC et qui a le plus réduit le BIC huit fois sur dix.

### 4.2.2. Compétition

Les indices de compétition retenus diffèrent en fonction des strates et espèces. Au niveau du peuplement deux indices montrent des différences moyennes entre arbres morts et arbres vivant (g10 : p-value = 0,04532 et g15 : p-value = 0,00054), l'indice g15 à largement été retenu par l'AIC et le BIC. Pour la strate de canopée, *Knema attenuata* et *Humboldtia brumonis* qui ne présentent pas de différences d'indice moyen entre arbres morts et vivants, c'est l'indice sélectionné de niveau peuplement (g15) qui a été arbitrairement retenu pour être testé dans les modèles. Pour les autres strates ou espèces, le tableau 6 présente l'indice qui a été retenu sur les critères de différence de moyenne entre arbres mort et vivants et sur les critères AIC et BIC.

	Strate émergente	Strate sous canopée	Strate sous bois	Vateria indica	Myristica dactyloïdes	Palaquium ellipticum
indice	G15surDBH2	G15surDBH	G15	G15surDBH2	g15surDBH	g15surDBH

Tableau 6 : indices de compétition à tester dans les modèles

### 4.2.3. Pente et exposition

La pente et l'exposition ont été retenues sous forme qualitative. La pente sera soit qualifiée de pente forte (>50%), soit de pente faible (<50%) pour permettre de tester plus facilement des éventuelles interactions avec des variables continues comme le diamètre ou l'accroissement. Au vu des distributions des expositions (figure 3), deux catégories d'exposition (exposition Est ou Ouest) ont été choisies pour refléter l'opposition de versant observée sur le terrain et de même pouvoir tester d'éventuelles interactions.

### 4.3. Résultats des modèles logistiques

Le tableau 7 présente les variables explicatives sélectionnées dans les modèles linéaires généralisés et dans les modèles linéaires généralisés mixtes. Le détail des coefficients et des statistiques de test sont présentés en annexe 3 pour les GLM et en annexe 6 pour les GLMM.

Pour tous les modèles mixtes proposés ici, les deux effets aléatoires ont été conservés systématiquement. Un modèle sans effet aléatoire mais ajusté sur l'ensemble des données a servi de point de comparaison pour évaluer le gain de vraisemblance apporté par les effets aléatoires. Remarquons qu'aucune interaction n'apparaît dans ces modèles.

Modèle		Variables explicatives							
		incgr	Incgr0.33	dbh	dbh2	pente	Compétition	date	arbre
Strate émergente	GLM	V	V			V			
	GLMM	V	V	V	V	V		V	V
Strate de canopée	GLM	V				V	g15		
	GLMM	V	V			V	g15	V	V
Strate de sous canopée	GLM	V	V				G15surDBH		
	GLMM	V	V				G15surDBH	V	V
Strate de sous bois	GLM		V	V			G15		
	GLMM	V	V	V	V		G15	V	V
Vateria indica	GLM		V	V	V	V	G15surDBH2		
	GLMM	V	V	V	V	V	G15surDBH2	V	V
Humboldtia brumonis	GLM	V							
	GLMM	V	V				g15	V	V
Myristica dactyloïdes	GLM	V	V				g15surDBH		
	GLMM	V	V	V	V			V	V
Knema attenuata	GLM	V							
	GLMM	V	V					V	V
Palaquium ellipticum	GLM		V			V	g15surDBH		
	GLMM	V		V		V		V	V

**Tableau 7 : modèles logistiques (GLM) et logistiques mixtes (GLMM) sélectionnés. V représente que la variable a été sélectionnée : elle permet une baisse des critères pénalisés AIC et BIC des modèles et le coefficient associé est significativement différent de zéro.**

Les courbes ROC associées aux GLM sont présentées en annexe 5 et le tableau 8 récapitule les valeurs d'AUC en fonction de l'échelle d'observation et du modèle.

		Strate émergente	Strate de canopée	Strate de sous canopée	Strate de sous bois	Vateria indica	Humboldtia brumonis	Myristica dactyloïdes	Knema attenuata	Palaquium ellipticum
AUC	GLM	0,7591	0,6947	0,7483	0,6367	0,7679	0,6245	0,7247	0,7045	0,7763
	GLMM	0,7702	0,6856	0,7834	0,6543	0,7912	0,6345	0,7019	0,6947	0,8024

**Tableau 8 : Valeurs des AUC (Area Under Curve) des analyses ROC**

La figure montre 5 la forme du modèle obtenu pour *Vateria indica* en l'absence d'effet aléatoire arbre ou date. Les prédictions obtenues par les modèles mixtes correspondants sont très proche, on observe d'ailleurs que les coefficients estimés par les deux types de modèles sont très voisins (tableau 9).

	Modèle à effets fixes	Modèle à effet aléatoire
Intercept	-0.4701835	-3.7404
incgr	/	2.2437530
incgr0.33	-2.3358523	-3.09821
SlopeFactforte	0.7902735	0.3796496
dbh	-0.0388989	-0.03061
dbh2	0.0004182	0.0003102
G15surDHB2	-0.0241508	-0.003192

Tableau 9 : coefficient des modèles associés à *Vateria indica*

La figure 5 montre que la probabilité de mort est diminuée avec l'accroissement. Pour les faibles valeurs d'accroissement, on observe une forme en U en fonction du diamètre. Dans le cas de pente forte, pour des valeurs d'accroissement et de diamètre similaires, on observe une probabilité de mort plus importante que dans les pentes faibles. Dans cette situation, on observe l'influence de l'accroissement et celui de la pente mais l'effet de la compétition reste très peu visible.

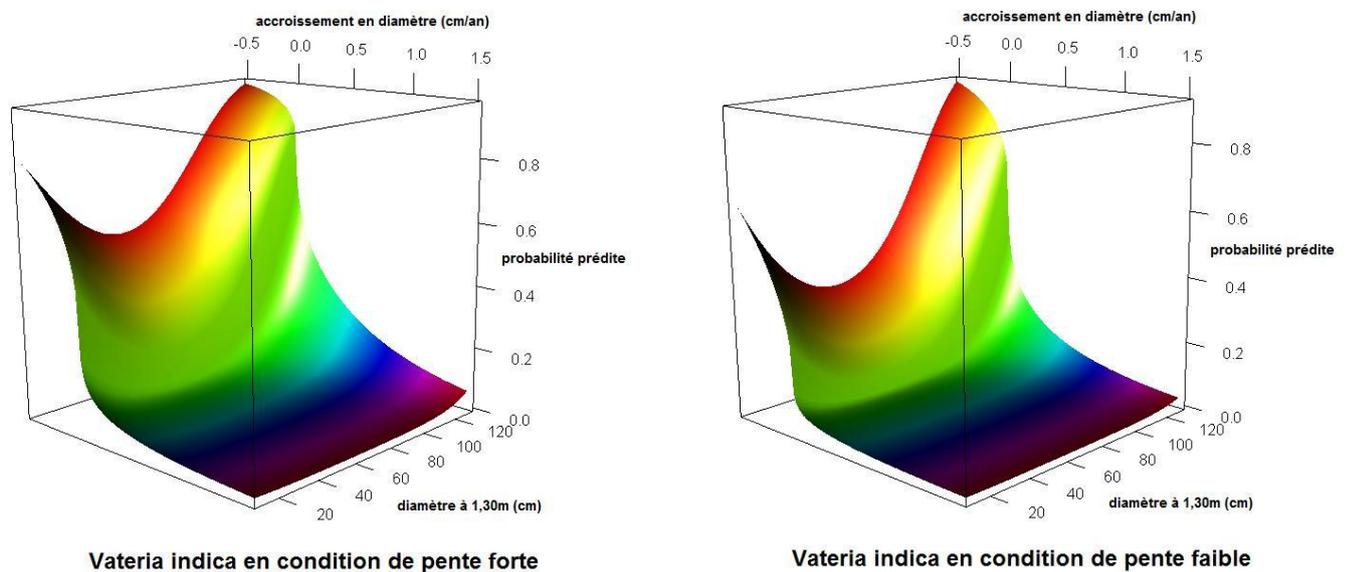
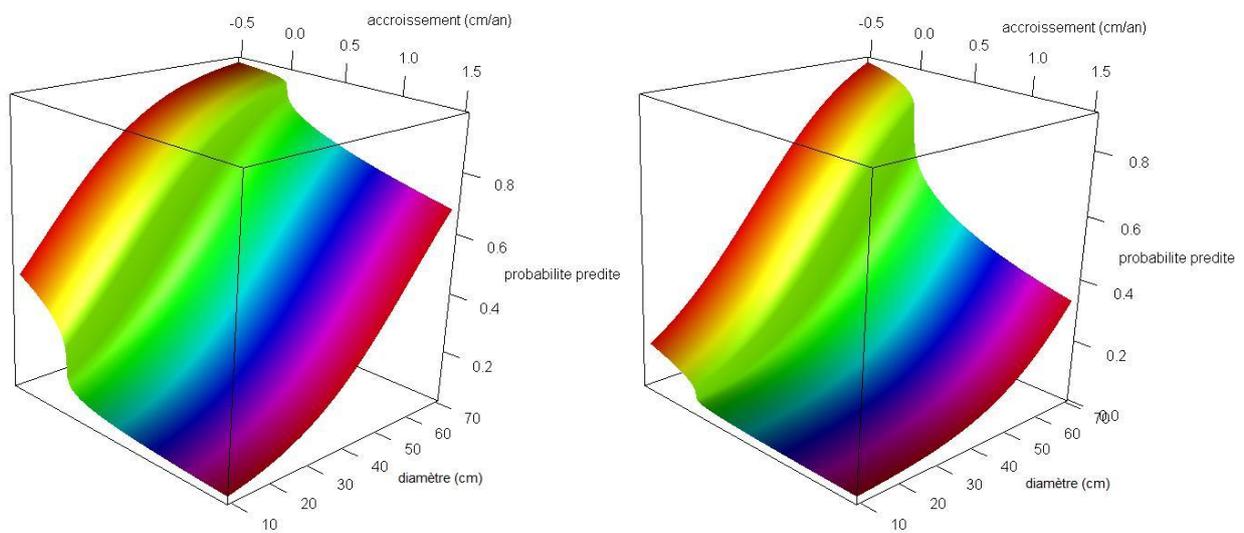


Figure 5 : probabilités prédites par le modèle sans effet aléatoire pour *Vateria indica* en fonction de l'accroissement et du diamètre et de la pente (la compétition a été fixée à sa valeur moyenne : g15surDBH2 = 0.2059).

Pour *Palaquium ellipticum* (figure 6) : la probabilité de mort prédite diminue avec l'accroissement. On n'observe pas de profil en U mais une probabilité plus importante pour les forts diamètres quelle que soit la valeur de l'accroissement.

Pour la strate émergente et la strate de canopée, les tendances sont les mêmes que celles identifiées pour *Vateria indica* et *Palaquium ellipticum* qui représentent ces strates.

Pour *Humboldtia brumonis*, *Knema attenuata* et *Myristica dactyloïdes*, seule la relation entre l'accroissement et la mortalité est nette. Les mêmes tendances sont observées pour la strate de sous canopée et la strate de sous bois.



*Palaquium ellipticum* en condition de forte pente

*Palaquium ellipticum* en condition de faible pente

**Figure 6 : probabilités prédites par le modèle sans effet aléatoire pour *Palaquium ellipticum* en fonction de l'accroissement et du diamètre et de la pente (la compétition a été fixée à sa valeur moyenne :  $g_{15surDBH2} = 3.298$ ).**

Le test d'ajustement de Hosmer et Lemeshow rejette systématiquement l'hypothèse nulle de bon ajustement du modèle. Ce test étant reconnu peu puissant, les approches graphiques ont donc été préférées pour interpréter l'ajustement aux données.

## 5. Discussion

---

Le premier constat est que les modèles qui sont ici fournis n'ont pas un pouvoir prédictif très fort. En effet, les distributions de probabilités prédites sont parfois quasiment confondues pour les arbres morts et les arbres vivants. Les analyses ROC montrent par exemple que les modèles proposés pour *Humboldtia brumonis* et la strate de sous-bois en général sont mauvais en termes de prédictions. En revanche, il faut tout de même remarquer que les modèles proposés pour *Vateria indica* et *Palaquium ellipticum* ainsi que pour la strate émergente et la strate de canopée, ont un pouvoir prédictif relativement meilleur.

### 5.1.1. Accroissement, diamètre et compétition

L'accroissement est, quelle que soit l'échelle d'observation (strate ou espèce), la meilleure variable prédictive de la mortalité (variable dont l'ajout provoque la plus forte hausse de vraisemblance des modèles). Néanmoins, cette variable est complétée dans presque tous les exemples par d'autres variables explicatives (sauf pour *Humboldtia brumonis* où il s'agit de la seule variable significativement explicative). Notons d'ailleurs que les modèles proposés pour *Humboldtia brumonis* sont ceux dont l'ajustement est le plus mauvais (tableau 8). Bien que l'accroissement semble une variable synthétique (de la compétition, de la taille, de la vigueur en générale) et apparaisse dans les modèles comme le meilleur prédicteur de la mortalité, des variables qui pourraient paraître redondantes apportent tout de même une part d'explication dans les modèles.

Pour les strates 1, 2 et 3 (canopée, sous canopée et sous bois) ainsi que pour les espèces testées (à l'exception de *Humboldtia brumonis* qui représente la strate de sous bois), le diamètre et / ou l'indice de compétition sont aussi des variables explicatives significatives. Plusieurs mécanismes sont retranscrits par l'accroissement, mais il est très probable qu'ils n'agissent pas de la même façon sur cette variable et que la manière dont l'accroissement apparaît dans le modèle traduise mal les différentes relations.

La part d'explication apportée par les indices de compétition reste tout de même faible, ce qui est peut être dû à la simplicité de ces indices utilisés pour traduire un phénomène complexe. Olono et al, (2009) ont montré qu'en plus de la densité et du diamètre des voisins, l'identité (appartenance à une espèce) avait de l'importance en termes de compétition dans le cas d'une forêt à faible mélange. Il serait intéressant de tester cette hypothèse pour les forêts fortement diversifiées. Les auteurs proposent de pondérer la surface terrière des voisins par un

coefficient propre à leur compétitivité spécifique. Dans le cas d'Uppangala, il serait, dans un premier temps, possible d'attribuer un coefficient de compétitivité aux espèces en se basant sur leur appartenance à une strate.

L'exemple de *Vateria indica* (figure 5) fait apparaître un profil de mortalité en U avec le diamètre. La mortalité juvénile est forte et tends à décroître avec la taille, ce phénomène est en général associé à la compétition forte qui existe en sous bois quand l'arbre n'a pas atteint la canopée. Une fois la canopée atteinte, la compétition pour la lumière est bien moins forte et la hausse du taux de mortalité alors observée peut être associée à la sénescence des arbres atteignant leurs dimensions maximales.

Dans le cas de *Palaquium ellipticum*, la mortalité est croissante avec le diamètre. Il peut aussi s'agir du même mécanisme de sénescence. Pour cette espèce, on n'observe pas de mortalité décroissante avec le diamètre (pour les faibles diamètres). Cela ne permet pas de dire que les juvéniles de l'espèce ne subissent pas de fortes pressions de compétition dans le jeune âge, il est possible que cette phase ne soit juste pas observée (les arbres sont recrutés pour les mesures à partir de 10 cm de diamètre).

### 5.1.2. Pente

Les modèles indiquent que la pente est un facteur qui augmente significativement la mortalité pour les strates hautes (émergente et canopée) ainsi que pour les deux espèces qui représentent ces strates (*Vateria indica* et *Palaquium ellipticum*). Ces espèces qui atteignent de grandes dimensions sont en effet susceptibles de subir plus lourdement les aléas climatiques de type tempête (Durrieu de Madron, 1993). Ces espèces susceptibles au chablis présentent alors une mortalité plus importante sur pente forte, leur grande taille et l'éventuel déséquilibre de houppier provoqué par la situation de pente (Moravie 1999) contribue à expliquer ce phénomène d'instabilité.

La pente ne semble en revanche pas augmenter la mortalité pour les espèces des strates inférieures, le chablis d'un arbre dominant entraîne pourtant en générale des chablis secondaires. On observe sur la figure 5 que l'effet de la pente sur la probabilité de mort de l'arbre se manifeste pour les accroissements faibles. Cela peut signifier que les arbres les moins vigoureux sont plus touchés par les chablis ou que le chablis n'est pas le seul mécanisme de mortalité sur lequel la pente a une influence. Il est par exemple envisageable que dans les situations de forte pente, les conditions de sol soient moins bonnes (sol plus lessivé et donc moins riche, sol plus fin et réserve en eau plus faible, etc) et provoquent

une sénescence pour des diamètres plus faibles. Par des mécanismes d'affaiblissement de la vigueur des arbres, la pente pourrait aussi par exemple entraîner une plus forte susceptibilité à des attaques de pathogènes. En effet, les sols de la forêt d'Uppangala sont pauvres et la fertilité est principalement due à la litière qui est peut être moyennement abondante en octobre mais très irrégulièrement répartie en raison de la pente et des blocs rocheux d'après Ferry (1994). La pente peut donc aussi être un facteur qui accroît la mortalité sur pied des arbres, entraînant ainsi à priori moins de dégâts parmi les strates inférieures. Il est donc difficile de dire si la pente influence la mortalité par le biais des chablis ou d'autres mécanismes.

Dans le cas de *Palaquium ellipticum*, il semble que la mortalité soit moins liée à une faible vigueur (figure 6) en effet, pour les gros diamètres, la mortalité est systématiquement prédite plus forte, même pour de forts accroissements. Sur pente faible, la mortalité reste associée majoritairement à de faibles accroissements et donc à priori des causes de mortalité régulières, alors que sur pente forte, la relation entre l'accroissement est beaucoup moindre. En effet, sur pente faible et pour un gros diamètre, la probabilité de mort est plus de deux fois plus importante pour de forts accroissements que pour de faibles accroissements. Sur pente forte, la différence est bien plus faible (figure 6). Il semble donc que cette espèce soit plus touchée par des évènements de mortalité irrégulière qui n'entraîne pas de baisse de vigueur avant la mort. Le chablis semble cette fois un bon candidat, mais n'est vraisemblablement pas la seule causes de mortalité associée à la taille de l'arbre.

Pour ces deux espèces présentes dans les strates supérieures, la pente semble donc avoir une forte influence sur les prédictions de mortalité. Néanmoins il apparaît que les mécanismes en jeu pour les deux espèces principales ne soient pas les mêmes. La mortalité de *Vateria indica* semble très fortement liée à un faible accroissement, traduisant plutôt une mortalité de type régulière, alors que les modèles prédisent une forte mortalité même pour les forts accroissements pour *Palaquium ellipticum*, traduisant que la mortalité de type irrégulière pourrait avoir plus d'importance pour *Palaquium ellipticum*.

### **5.1.3. Apport du modèle mixte**

Les modèles mixtes ont améliorés l'ajustement des modèles et même si leur capacité prédictive n'a pas été améliorée beaucoup améliorée, ces modèles seront plus proches des mécanismes réels dans un cadre de simulation de dynamique en permettant de mieux

structurer la variabilité (variabilité temporelle et intraspécifique) qui n'était pas expliquée par les modèles à effets fixes.

#### **5.1.4. Perspectives d'amélioration**

Différentes raisons peuvent être à l'origine des difficultés rencontrées pour bien prédire ce phénomène de mortalité. Il s'agit d'une part d'un évènement rare et dont on observe qu'une fois la réalisation par individu. Un jeu de données plus conséquent aurait probablement permis de mieux identifier le rôle des différentes variables explicatives. Néanmoins, l'amélioration que cela apporterait resterait sûrement limitée à cause de la nature même de la mortalité. En effet, la mortalité est associée à différentes causes qui impliquent potentiellement différentes relations entre le phénomène de mortalité et les variables explicatives. On a pourtant ici cherché à modéliser en une seule fois les différents mécanismes entraînant la mortalité (compétition, sénescence, chablis). Cela vient du fait que les données ne permettent pas de d'identifier les différents types de mortalités. Il serait d'ailleurs intéressant à l'avenir de réaliser une identification systématique du type de mortalité lors des recensements annuels. La connaissance du type de mortalité (notamment, mortalité sur pied ou chablis) associée à chaque évènement de mort pourrait permettre de modéliser les phénomènes séparément (au moins de distinguer mortalité régulière et irrégulière)

De plus, certains évènements qui causent potentiellement la mortalité sont mal ou pas du tout prédits (attaque d'insecte, tempête ou sécheresse exceptionnelle, etc.), il reste donc une part aléatoire qu'il est difficile de prédire.

Enfin, il est difficile de connaître le moment exact de la mort. La mortalité a pu être enregistrée un an trop tôt ou un an trop tard par exemple. Les caractéristiques associées à l'arbre au moment de l'enregistrement de la mort peuvent différer des réelles caractéristiques du moment de la mort (notamment en termes d'accroissement). Les erreurs de mesure introduisent donc une certaine imprécision qu'il serait d'ailleurs intéressant de quantifier et pourrait permettre de spécifier la structure des résidus dans les modèles.

Le manque de performance des modèles proposés peut venir d'une mauvaise forme des modèles. Dans ce cas, des méthodes semi ou non-paramétriques de modélisation permettant de s'affranchir des hypothèses sur les formes de modèles (Vieilledent et al., 2009 ; Wyckoff and Clark, 2000) seraient peut être plus adaptés pour modéliser la mortalité. Il serait notamment intéressant de mettre en œuvre ces méthodes et d'évaluer le gain apporté aux modèles.

# Références

---

- Agresti, A., 2002. *Categorical data analysis*, New York, Chichester, Brisbane, Toronto, Singapore: Wiley -Interscience Publication.
- Agresti, Alan. 1996. *An Introduction to Categorical Data Analysis*. John Wiley and Sons, Inc.
- Akaike, H., 1974. A new look at statistical model identification. *IEEE Transactions on Automatic Control* AU-19: 716-722
- Bolker, Benjamin M, Mollie E Brooks, Connie J Clark, Shane W Geange, John R Poulsen, M Henry H Stevens, and Jada-Simone S White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution* 24, no. 3 (March): 127-35.
- Canham C.D., Papaik M.J., and Latty E.F., 2001. Interspecific variation in susceptibility to windthrow as a function of tree size and storm severity for northern temperate tree species. *Canadian Journal of Forest Research*. 31: 1–10.
- Casper, B.B. & Jackson, R.B. (1997) Plant competition underground. *Annual Review of Ecology and Systematics*, 28, 545–570.
- Chao, K. J. et al. 2008. Growth and wood density predict tree mortality in Amazon forests. *Journal of Ecology* 96: 281–292.
- Chao, Kuo-Jung, Oliver L. Phillips, Emanuel Gloor, Abel Monteagudo, Armando Torres-Lezama, and Rodolfo Vásquez Martínez. 2008. Growth and wood density predict tree mortality in Amazon forests. *Journal of Ecology* 96, no. 2 (March): 281-292.
- Clark, J. S., D. Bell, C. Chu, B. Courbaud, M. Dietze, M. Hersh, J. HilleRisLambers, I. Ibanez, S. LaDeau, S. McMahon, J. Metcalf, J. Mohan, E. Moran, L. Pangle, S. Pearson, C. Salk, Z. Shen, D. Valle, and P. Wyckoff., 2010. High-dimensional coexistence based on individual variation: a synthesis of evidence. *Ecological Monographs* 80 : 569-608.
- Clark, D. A., and D. B. Clark. 1999. Assessing the growth of tropical rain forest trees: issues for forest modeling and management. *Ecological Applications* 9:981-997.
- Clark, J.S., Dietze, M., Chakraborty, S., Agarwal, P., Ibáñez, I., LaDeau, S., and Wolosin, M., 2007. Resolving the biodiversity paradox. *Ecology Letters* 10 : 647-662.
- Clark, James S. 2004. Why environmental scientists are becoming Bayesians. *Ecology Letters* 8, no. 1 (December): 2-14.
- Collett, D., 2003. *Modelling Survival Data in Medical Research, Second Edition*. Boca Raton: Chapman & Hall/CRC.
- Coomes DA, Duncan RP, Allen RB, Truscott J (2003) Disturbances prevent stem size-density distributions in natural forests from following scaling relationships. *Ecology Letters* 6: 980–989.
- Coomes, David A., and Robert B. Allen. 2007. Mortality and tree-size distributions in natural mixed-age forests. *Journal of Ecology* 95, no. 1 (January): 27-40.
- Delbart, N, P Ciais, J Chave, N Viovy, Y Malhi, and T Le Toan. 2010. Mortality as a key driver of the spatial distribution of aboveground biomass in Amazonian forest : results from a dynamic vegetation model. *Biogeosciences*: 3027-3039.
- Diggle, P.J., Liang, K.-Y., Zeger, S.L., 1994. *Analysis of longitudinal data*. Oxford, Clarendon Press, 253 p.
- Durrieu de Madron, L. 1993. Mortalité, chablis et rôle des trouées dans la sylvigénèse avant et après exploitation sur le dispositif d'étude sylvicole de Paracou - Guyane française. Thèse de doctorat.
- Elouard, C., R. Pélissier, F. Houllier, J.-P. Pascal, M. Durand, S. Aravajy, M. A. Moravie, C. Gimaret-Carpentier, and B. R. Ramesh. 1997b. Monitoring the structure and dynamics of a dense moist evergreen forest in the Western Ghats (Kodagu District, Karnataka, India). *Tropical Ecology* 38:193-214.
- Ferry, B. 1994. Etude des humus forestiers des Ghâts occidentaux en Inde dusud - facteurs climatiques, édaphiques et biologiques intervenant dans le stockage de la matière organique du sol. Publication du département d'écologie No. 34, Institut Français de Pondichéry, Inde. 260pp.

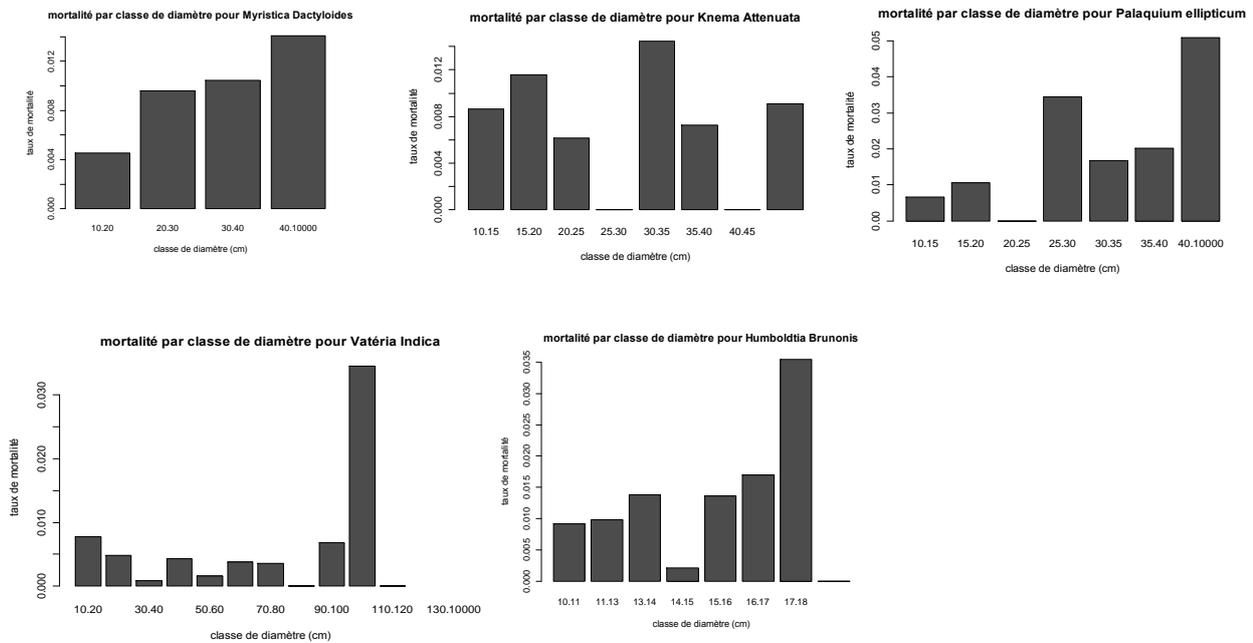
- Gower S.T.,McMurtrie R.E., andMurty D., 1996. Aboveground net primary production decline with stand age: Potential causes. *Trends Ecology and Evolution*. 11: 378–382.
- Hadfield, J. D., 2010.MCMC methods for Multi response Generalised Linear Mixed Models : The MCMCglimm R Package. *Journal of Statistical Software*, 33(2),1 : 22
- Hamilton, D. A., 1974. Event probabilities estimated by regression.USDA For. Serv. Res. Pap. INT-152.
- Hamilton, D. A., 1986. A logistic model of mortality in thinned and unthinned mixed conifer stands of northern Idaho. *Forest Science* 32:989–1000.
- Hamilton, D. A., Edwards, B. M.,1976. Modeling the probability of individual tree mortality. U.S.D.A. For. Serv. Res. Pap. INT-185
- Hawkes, Corinna. 2000. Woody plant mortality algorithms: description, problems and progress. *Ecological Modelling* 126, no. 2-3: 225–248.
- Hosmer, D. W., Lemeshow, S., Klar, J.,1988. Goodness-of-fit testing for multiple logistic regression analysis when the estimated probabilities are small. *Biometrical Journal*. 30,911–924.
- Hosmer, D.W., Lemeshow, S., 1989. *Applied Logistic Regression*. Wiley, New York, p. 309.
- Hosmer, DW, T Hosmer, S Le Cessie, and S Lemeshow. 1997. A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in medicine* 16, no. 9 (May 15): 965-80.
- Hubbard R.M., Bond B.J., and RyanM.G., 1999. Evidence that hydraulic conductance limits photosynthesis in old Pinus ponderosa trees. *Tree Physiol*. 19: 165–172.
- Kunstler, Georges, David a. Coomes, and Charles D. Canham. 2009. Size-dependence of growth and mortality influence the shade tolerance of trees in a lowland temperate rain forest. *Journal of Ecology* 97, no. 4 (July): 685-695.
- Kurz WA, Dymond CC, Stinson G, Rampley GJ, Neilson ET, et al. (2008) Mountain pine beetle and forest carbon feedback to climate change. *Nature* 452: 987–990.
- Lee, Y.J., 1971. Predicting mortality for even-aged stands of lodgepole pine. *For. Chron*. 47: 29–32.
- Lee, Y., J.A. Nelder, and Y. Pawitan. 2006. Generalized linear models with random effects: unified analysis via H-likelihood. Analysis. CRC Press.
- Lines, Emily R, David a. Coomes, and Drew W Purves. 2010. Influences of forest structure, climate and species composition on tree mortality across the eastern US. *PloS one* 5, no. 10 (January) ;
- Link, William a, and Richard J Barker. 2006. Model weights and the foundations of multimodel inference. *Ecology* 87, no. 10 (October): 2626-35.
- McCullagh P. and Nelder, J. A. (1989) *Generalized Linear Models*. London: Chapman and Hall
- Madelaine, Cécile. 2009. Dynamique des peuplements forestiers tropicaux hétérogènes : variabilité inter et intraspécifique de la croissance des arbres et trajectoires de développement en forêt dense humide sempervirente, dans les Ghâts occidentaux de l’Inde. *Thèse*. Université Montpellier 2.
- Manion, P. D., 1981. *Tree Disease Concepts*. Englewood Cliffs, Prentice-Hall, p.399.
- Monserud, R. A., 1976. Simulation of forest tree mortality. *Forest Science*. 22:438–444.
- Monserud, R.A., and Sterba, H. 1999. Modeling individual tree mortality for Austrian forest species. *Forest Ecology and Management*. 113(2– 3): 109–123
- Moravie, M.A., 1996. Analyse de la mortalité naturelle en forêt tropicale humide des Ghâts Occidentaux de l’Inde. Rapport de mission à l’étranger. ENGREF, Montpellier
- Moravie, M. A. 1999. Un modèle arbre dépendant des distances pour l’étude des relations entre la dynamique et la structure spatiale d’une forêt dense humide sempervirente. Ph.D. Thesis. Université Claude Bernard - LYON I, Lyon.
- Muller-Landau HC, Condit RS, Harms KE, Marks CO, Thomas SC, et al. (2006) Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. *Ecology Letters* 9: 589–602.
- Pascal, J.-P. 1988. Wet evergreen forests of the Western Ghats of India : ecology, structure, floristic composition and succession. *Travaux de la Section Scientifique et Technique*, Tome 20 bis. French Institute of Pondicherry, India.
- Pascal, J.-P., and R. Pélissier. 1996. Structure and floristic composition of a tropical evergreen forest in southwest India. *Journal of Tropical Ecology* 12: 195-218.
- Peet R.K. and Christensen N.L., 1987. Competition and tree death. *BioScience* 37: 586–595.

- Pelissier, Raphael, Jean-pierre Pascal, N. Ayyappan, B. R. Ramesh, S Aravajy, and S. Ramalingam. 2011. Twenty years tree demography in an undisturbed Dipterocarp Permanent Sample Plot at Uppangala, Western Ghâts of India. *Ecology*: 1-35.
- Pretzsch, Hans, and Peter Biber. 2010. Size-symmetric versus size-asymmetric competition and growth partitioning among trees in forest stands along an ecological gradient in central Europe. *Canadian Journal of Forest Research* 40, no. 2: 370-384.
- Purves, D. and Pacala, S. 2008. Predictive models of forest dynamics. *Science* 320: 1452–1453.
- R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Ramesh, B. R., Ayyappan, N., Grard, P., Prosperi, J., Aravajy, S. & Pascal, J.-P., 2010. BIOTIK: Biodiversity Informatics and co-Operation in Taxonomy for Interactive Shared Knowledge Base. Western Ghats v.1.0. A multimedia identification system of evergreen tree species of the Western Ghats, India. Collection Ecologie n°46, French Institute of Pondicherry. DVD (<http://www.biotik.org/>).
- Reineke, L.H. 1933. Perfecting a stand-density index for even-aged forests. *Journal of Agricultural Research* 46:627-638.
- Ripley B.D. 1977. Modelling spatial patterns. *Journal of the Royal Statistical Society B*, 39:172-192.
- Rüger, Nadja, Andreas Huth, Stephen P. Hubbell, and Richard Condit. 2011. Determinants of mortality across a tropical lowland rainforest community. *Oikos* : 1-10
- Salaün, P. 1995. Représentations, utilisations et transformations de la richesse spécifique dans quatre communautés forestières des Ghâts occidentaux. Thèse de Doctorat, Université Pierre et Marie Curie, Paris. 237 pp.
- Schwarz, Gideon E. (1978). Estimating the dimension of a model. *Annals of Statistics* 6 (2): 461–464.
- Schtickzelle, M., 1981. Pierre-François Verhulst (1804-1849). La première découverte de la fonction logistique, *Population* 3, p. 541-556.
- Shugart, H.H., 1984. A Theory of Forest Dynamics: The Ecological Implications of Forest Succession Models. Springer-Verlag, New York, p. 278.
- van Mantgem PJ, Stephenson NL (2007) Apparent climatically induced increase of tree mortality rates in a temperate forest. *Ecology Letters* 10: 909–916.
- Verhulst, Pierre-François, 1838. Notice sur la loi que la population poursuit dans son accroissement, dans *Correspondance mathématique et physique*, n° 10, p. 113-121
- Vieilledent, Ghislain, Benoît Courbaud, Georges Kunstler, and Jean-François Dhôte. 2010. Mortality of silver fir and Norway Spruce in the Western Alps – a semi-parametric approach combining size-dependent and growth-dependent mortality. *Annals of Forest Science* 67, no. 3 (February): 305.
- Vieilledent, Ghislain, Benoît Courbaud, Georges Kunstler, Jean-François Dhôte, and James S. Clark. 2009. Biases in the estimation of size-dependent mortality models: advantages of a semiparametric approach. *Canadian Journal of Forest Research* 39, no. 8: 1430-1443.
- Wunder, J. et al. 2008. Predicting tree mortality from growth data: how virtual ecologists can help real ecologists. *Journal of Ecology* 96: 174–187.
- Wyckoff, P.H., and James S. Clark. 2000. Predicting tree mortality from diameter growth: a comparison of maximum likelihood and Bayesian approaches. *Canadian Journal of Forest Research* 30, no. 1: 156–167.
- Wyckoff P.H. and Clark J.S., 2002. The relationship between growth and mortality for seven co-occurring tree species in the southern Appalachian Mountains. *J. Ecol.* 90: 604–615.
- Yang, Yuqing, Stephen J. Titus, and Shongming Huang., 2003. Modeling individual tree mortality for white spruce in Alberta. *Ecological Modelling* 163, no. 3 (May): 209-222.
- Yao, Xiaohong, Stephen J. Titus, and S. Ellen MacDonald. 2001. A generalized logistic model of individual tree mortality for aspen, white spruce, and lodgepole pine in Alberta mixed wood forests. *Canadian Journal of Forest Research* 31, no. 2: 283-291.
- Zens, M. 2003. Dealing with death data: individual hazards, mortality and bias. *Trends in Ecology & Evolution* 18, no. 7 (July): 366-373.
- Zuur, A.F., E.N. Ieno, N.J. Walker, A.A. Saveliev, and G.M. Smith. 2009. *Mixed effects models and extensions in ecology with R*. Genetics. Springer Verlag.



# Annexes

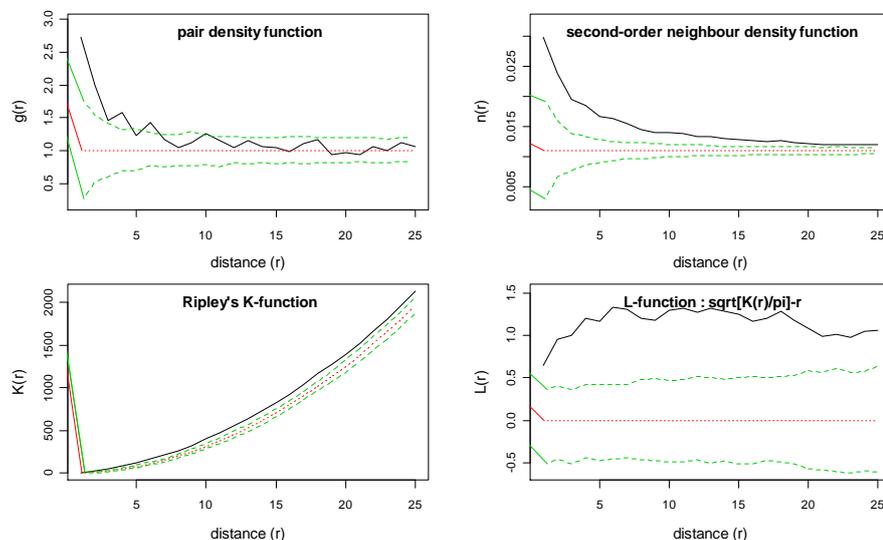
## Annexe 1 : mortalité observée par classe de diamètre sur l'ensemble de la période de 20 ans



## Annexe 2 bis : Analyse de Ripley sur la mortalité

analyses de Ripley pour les arbres morts depuis 20 ans (entre census0 et census5)

— obs      - - - - - theo (CSR)      - - - - - 99 % CI of CSR



## Annexe 2 : comparaison de moyennes des covariables en fonction du statut Mort ou Vivant

Comparaison de moyenne des variables potentiellement explicatives en fonction du statut mort ou vivant. Le tableau présente ici les probabilités critiques des tests de Student en fonction de la variable considérée et du niveau de description dans le peuplement

	niveau de description										
	peuplement	vain	hubr	myda	pael	knat	strate0	strate1	Strate2	strate3	
dbh	0,25081	0,35860	0,65310	0,06178	0,01338	0,66860	0,34652	0,46083	0,06763	0,01862	
dbh2	0,35002	0,95993	0,64256	0,08358	0,02374	0,71710	0,88493	0,67916	0,07777	0,02140	
slope	0,00008	0,04046	0,43484	0,40860	0,07587	0,09931	0,16715	0,03012	0,28775	0,01720	
aspect	0,57633	0,42852	0,26576	0,41933	0,94785	0,20125	0,49076	0,00837	0,17990	0,94982	
incgr	0,00000	0,00005	0,00003	0,34829	0,20159	0,00002	0,00000	0,00000	0,00228	0,00361	
incgr0,33	0,00000	0,00000	0,00000	0,00020	0,02648	0,00002	0,00000	0,00000	0,00000	0,00000	
expgr	0,00000	0,00000	0,00003	0,01235	0,00811	0,00002	0,00000	0,00000	0,00000	0,00312	
G5	0,50587	0,39427	0,23552	0,70798	0,27867	0,14447	0,18474	0,26233	0,74857	0,02000	
G10	0,30944	0,39572	0,87149	0,54380	0,00308	0,95734	0,12420	0,08803	0,50403	0,01397	
G15	0,05003	0,58003	0,67285	0,04124	0,01207	0,72016	0,30700	0,16199	0,05765	0,00139	
g5	0,20031	0,85528	0,23026	0,32198	0,37630	0,14995	0,92978	0,53517	0,28054	0,03796	
g10	0,04532	0,72625	0,87188	0,90899	0,13804	0,95079	0,66107	0,10896	0,99583	0,05095	
g15	0,00054	0,25888	0,65916	0,09374	0,27199	0,68049	0,55131	0,14973	0,24905	0,00600	
variables	G5surDBH	0,73480	0,05273	0,25623	0,82846	0,09314	0,10643	0,03005	0,36773	0,73547	0,01759
	G10surDBH	0,84232	0,02947	0,71215	0,19891	0,00317	0,54450	0,01091	0,10750	0,19791	0,00505
	G15surDBH	0,79277	0,03887	0,81085	0,00895	0,02511	0,69076	0,02093	0,29817	0,03968	0,00249
	g5surDBH	0,59632	0,09219	0,24863	0,93181	0,21231	0,09965	0,06038	0,42552	0,93030	0,01799
	g10surDBH	0,62955	0,04966	0,70277	0,21643	0,00384	0,54882	0,01813	0,09297	0,21520	0,00515
	g15surDBH	0,56488	0,06739	0,80547	0,00675	0,03027	0,65984	0,03554	0,28888	0,03404	0,00202
	G5surDBH2	0,88516	0,02347	0,26505	0,66598	0,10083	0,13453	0,01498	0,43417	0,48301	0,02270
	G10surDBH2	0,78068	0,01157	0,59066	0,25189	0,02015	0,45297	0,00474	0,16854	0,20486	0,00808
	G15surDBH2	0,64661	0,01697	0,90936	0,01911	0,07566	0,56756	0,00810	0,41658	0,06837	0,01189
	g5surDBH2	0,83795	0,02586	0,25691	0,72060	0,11028	0,13520	0,01878	0,44119	0,49348	0,01960
	g10surDBH2	0,85315	0,01319	0,57814	0,23869	0,01844	0,45072	0,00586	0,16071	0,18981	0,00652
	g15surDBH2	0,71293	0,01884	0,91341	0,01577	0,07155	0,55408	0,00948	0,41097	0,05800	0,00934

Les cases en bleu représentent les probabilités critiques inférieures à 5%

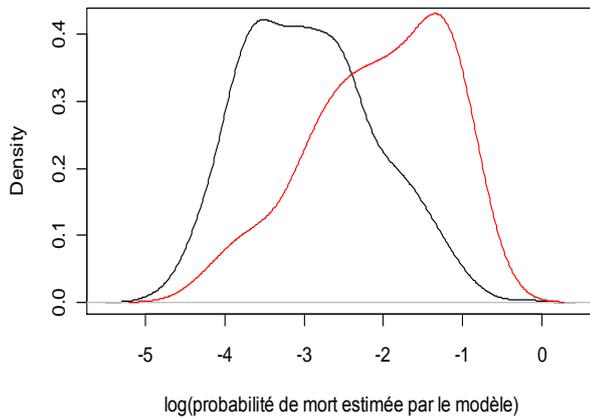
## Annexe 3 : GLM

données	variable	coef	Standard error	Z value	P value
<b>peuplement</b>	Intercept	-1.010755	0.224457	-4.503	6.70e-06 ***
	incgr0.33	-2.415895	0.225697	-10.704	< 2e-16 ***
	SlopeFactforte	0.303378	0.107768	2.815	0.00488 **
	incgr	1.357521	0.442114	3.071	0.00214 **
	g15	-0.019658	0.003579	-5.493	3.96e-08 ***
	PotStrata1	0.350376	0.170585	2.054	0.03998 *
	PotStrata2	0.103664	0.168782	0.614	0.53909
	PotStrata3	0.339285	0.158996	2.134	0.03285 *
<b>Strate 0 : émergente</b>	Intercept	-1.9667	0.2008	-9.792	< 2e-16 ***
	incgr	-1.8717	0.8697	2.152	0.0314 *
	incgr0.33	-3.1702	0.5256	-6.032	1.62e-09 ***
	SlopeFactforte	0.5693	0.2616	2.176	0.0295 *
<b>Strate 1 : canopée</b>	Intercept	-1.22654	0.24419	-5.023	5.09e-07 ***
	incgr0.33	-2.39030	0.30409	-7.860	3.83e-15 ***
	SlopeFactforte	0.51187	0.23304	2.197	0.0281 *
	g15	-0.17076	0.07321	-2.333	0.0197 *
<b>Strate 2 : sous canopée</b>	Intercept	-1.19246	0.19652	-6.068	1.30e-09 ***
	incgr0.33	-2.54596	0.42322	-6.016	1.79e-09 ***
	incgr	2.00679	0.76090	2.637	0.00835 **
	G15surDBH	-0.24324	0.08997	-2.703	0.00686 **
<b>Strate 3 : sous bois</b>	Intercept	-1.192715	0.338666	-3.522	0.000429 ***
	incgr0.33	-1.359287	0.264146	-5.146	2.66e-07 ***
	dbh	0.047625	0.015013	3.172	0.001513 **
	G15	-0.021906	0.005796	-3.779	0.000157 ***
<b>Vateria indica</b>	Intercept	-0.4701835	0.6290571	-0.747	0.4548
	incgr0.33	-2.3358523	0.4115303	-5.676	1.38e-08 ***
	SlopeFactforte	0.7902735	0.3141514	2.516	0.0119 *
	dbh	-0.0388989	0.0239467	-1.624	0.1043
	dbh2	0.0004182	0.0001973	2.120	0.0340 *
	G15surDHB2	-0.0241508	0.0105417	-2.291	0.0220 *
<b>Humboldtia brumonis</b>	Intercept	-1.6980	0.1349	-12.589	< 2e-16 ***
	incgr0.33	-1.6087	0.5003	-3.215	0.00130 **
<b>Myristica dactyloïdes</b>	Intercept	-0.14603	0.48438	-0.301	< 2e-16 ***
	incgr	3.97923	1.22966	3.236	0.0522 .
	incgr0.33	-3.26269	0.65061	-5.015	5.31e-07 ***
<b>Knema attenuata</b>	g15surDBH	-0.02960	0.01025	-2.889	9.2e-06 ***
	Intercept	-2.0343	0.2349	-8.662	< 2e-16 ***
	incgr0.33	-2.2859	0.7026	-3.253	0.00114 **
<b>Palaquium ellipticum</b>	Intercept	-0.6872	0.6276	-1.095	0.27352
	incgr0.33	-1.5985	0.6466	-2.472	0.01343 *
	SlopeFactforte	1.4058	0.5071	2.772	0.00556 **
	dbh	0.06708	0.01834	3.658	0.000255 ***

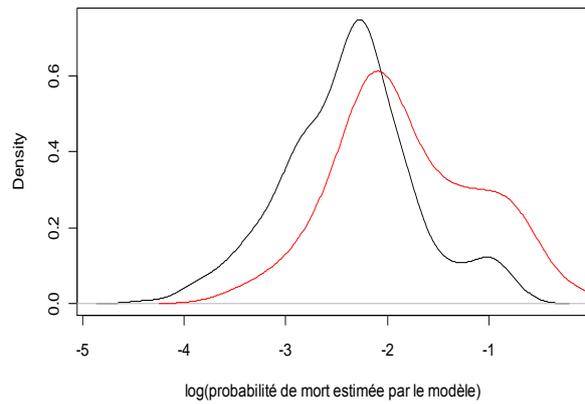
## Annexe 4 : Distributions des probabilités prédites par les GLM

En rouge : les arbres morts et noire : les arbres vivants.

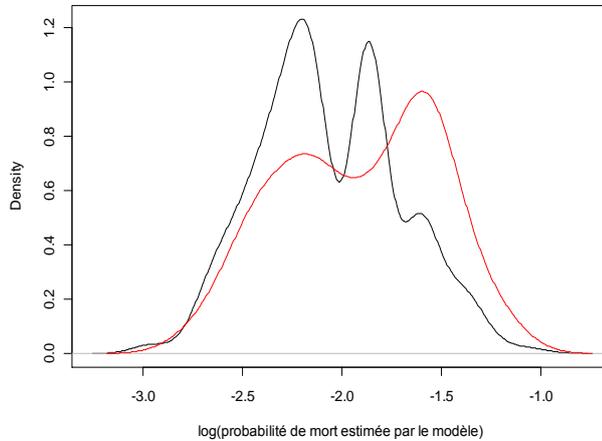
distribution des probabilités estimées pour *Vateria indica*



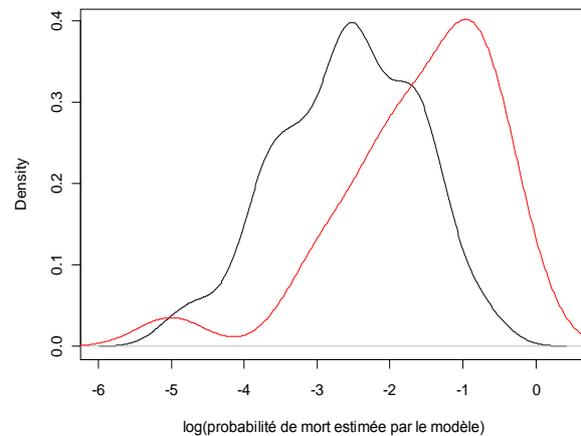
distribution des probabilités estimées pour *Mirystica dactyloides*



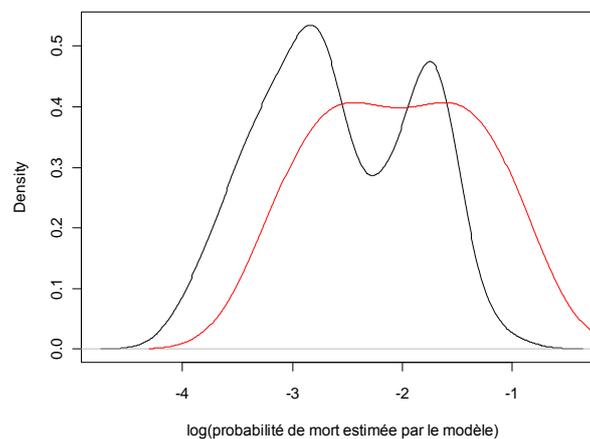
distribution des probabilités estimées pour *Humboldtia brumonis*



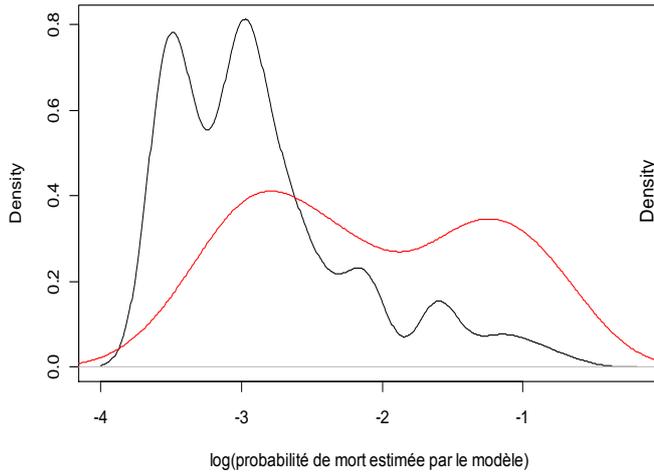
distribution des probabilités estimées pour *Pallaquium ellipticum*



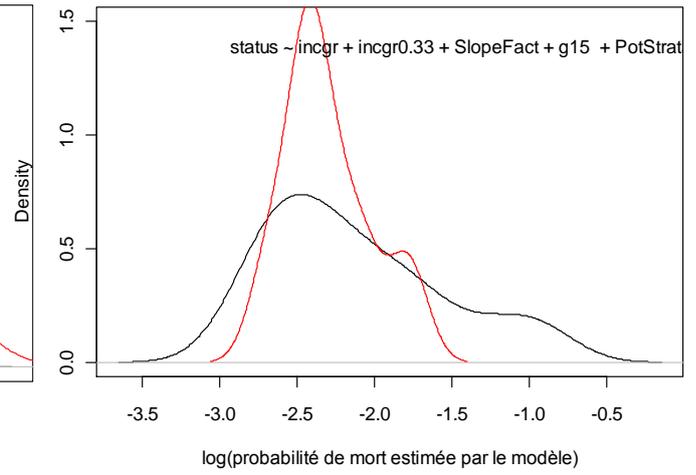
distribution des probabilités estimées pour *Knema attenuata*



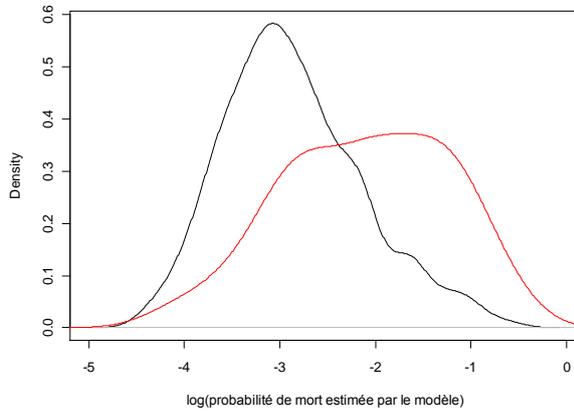
distribution des probabilités estimées pour la strate émergente



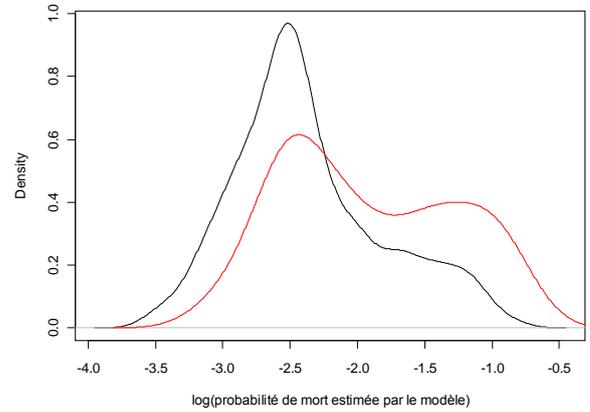
distribution des probabilités estimées pour tout le peuplement



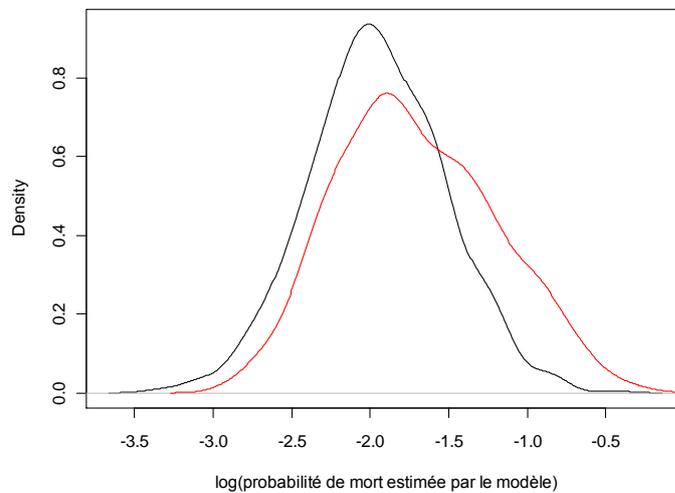
distribution des probabilités estimées pour la strate de canopée



distribution des probabilités estimées pour la strate de sous canopée

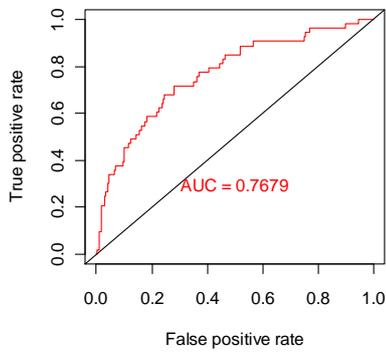


distribution des probabilités estimées pour la strate de sous bois

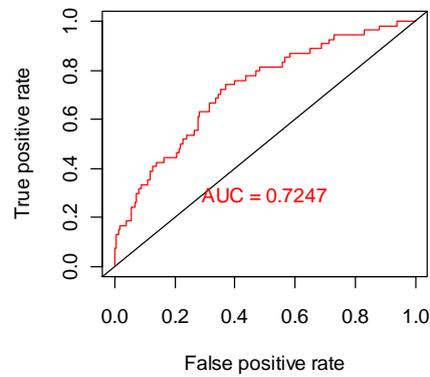


## Annexe 5 : Courbes ROC des GLM

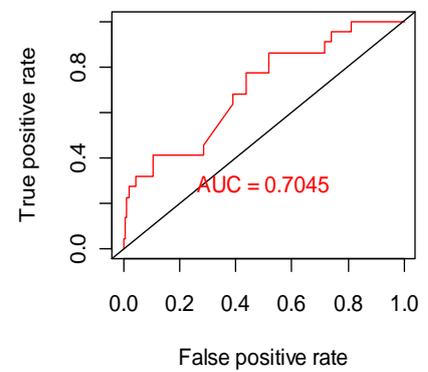
**courbe ROC : *Vateria indica***



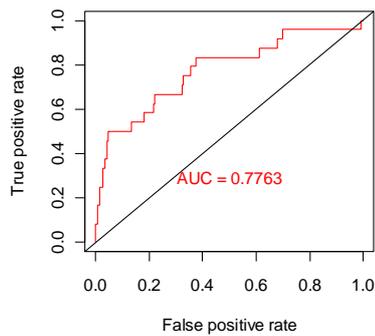
**courbe ROC : *Myristica dactyloides***



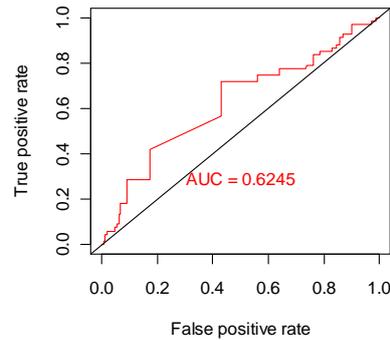
**courbe ROC : *Knema attenuata***



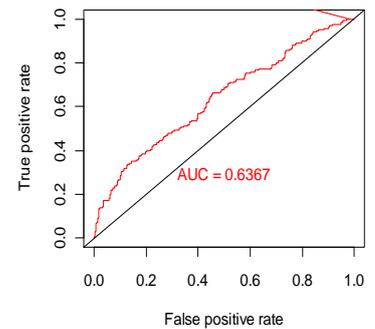
**courbe ROC : *Palaquium ellipticum***



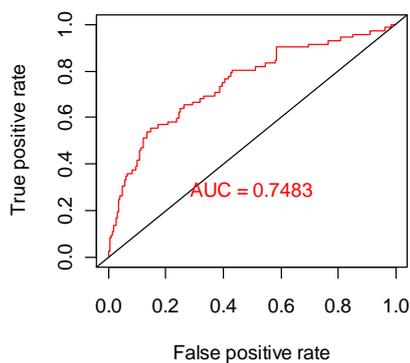
**courbe ROC : *Humboldtia brumonis***



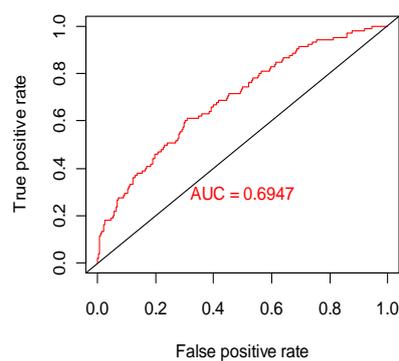
**courbe ROC : strate de sous bois**



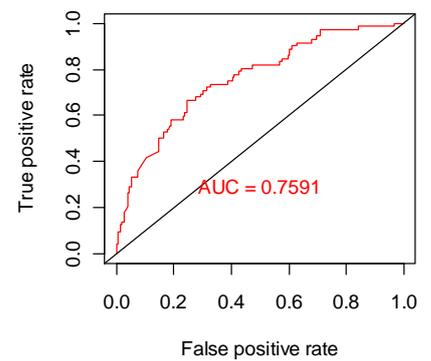
**courbe ROC : strate de canopée**



**courbe ROC : strate de sous canopée**



**courbe ROC : strate émergente**



## Annexe 6 : GLMM

données	variable	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
<b>peuplement</b>	(Intercept)	-4.293e+00	-4.472e+00	3.968e+00	4.102	<0.001 ***
	dbh	1.512e-02	5.770e-03	2.466e-02	2.224	<0.001 ***
	dbh2	-1.016e-04	-1.736e-04	-2.648e-05	5.190	0.002 **
	incgr	1.557e+00	1.277e+00	1.771e+00	7.705	<0.001 ***
	incgr0.33	-2.449e+00	-2.845e+00	-1.992e+00	3.813	<0.001 ***
	g15	-1.772e-03	-2.668e-03	-1.009e-03	5.362	<0.001 ***
	SlopeFactforte	2.347e-01	1.347e-01	3.333e-01	6.841	<0.001 ***
	PotStrata1	1.070e-01	-9.799e-02	3.308e-01	1.510	0.542
	PotStrata2	-8.813e-02	-2.150e-01	3.453e-02	4.699	0.196
	PotStrata3	1.638e-01	2.335e-02	3.756e-01	2.552	0.002 **
<b>Strate 0 : émergente</b>	Intercept	-3.969e+00	-4.202e+00	-3.763e+00	14.051	<0.001 ***
	dbh	-1.257e-02	-2.807e-02	3.446e-03	5.491	0.246
	dbh2	1.600e-04	2.571e-05	3.067e-04	6.079	0.018 *
	incgr	1.905e+00	7.373e-01	3.095e+00	3.487	<0.001 ***
	incgr0.33	-3.095e+00	-4.019e+00	-1.813e+00	3.251	<0.001 ***
<b>Strate 1 : canopée</b>	Intercept	-3.289517	-3.590427	-2.988792	3.924	<0.001 ***
	incgr	1.383028	0.649345	1.929160	7.028	<0.001 ***
	incgr0.33	-2.548234	-2.922117	-2.178679	6.731	<0.001 ***
	g15	-0.011724	-0.017124	-0.006547	2.979	<0.001 ***
	SlopeFactforte	0.163543	0.069836	0.253522	10.386	<0.001 ***
<b>Strate 2 : sous canopée</b>	Intercept	-3.308298	-3.500638	-3.076998	4.327	<0.001 ***
	incgr	1.697309	1.373243	2.043433	10.891	<0.001 ***
	incgr0.33	-2.015238	-2.524377	-1.504345	2.027	<0.001 ***
	G15surDBH	-0.010872	-0.016572	- 0.006263	1.749	<0.001 ***
<b>Strate 3 : sous bois</b>	Intercept	-4.347e+00	-4.767e+00	-3.987e+00	5.883	<0.001 ***
	dbh	8.136e-02	4.285e-02	1.202e-01	5.729	<0.001 ***
	dbh2	-9.368e-04	-1.877e-03	4.676e-06	5.150	0.018 *
	incgr	2.002e+00	1.324e+00	2.733e+00	6.706	<0.001 ***
	incgr0.33	-2.201e+00	-2.454e+00	-1.823e+00	5.816	<0.001 ***
	G15	-8.599e-03	-1.239e-02	-4.888e-03	3.830	<0.001 ***
<b>Vateria indica</b>	(Intercept)	-3.7404	-4.06718	-3.3919	6.154	<0.001 ***
	dbh	-0.03061	-0.04779	-0.01180	5.328	<0.001 ***
	dbh2	0.0003102	0.0001154	0.0005325	2.513	<0.001 ***
	incgr	2.2437530	1.4379190	3.0426590	4.724	<0.001 ***
	incgr0.33	-3.09821	-3.86380	-1.77071	2.918	<0.001 ***
	G15surDHB2	-0.003192	-0.00613	-0.00064	4.406	0.006 **
	SlopeFactforte	0.3796496	0.1279612	0.5909851	5.577	<0.001 ***
	(Intercept)	-3.756259	-3.848449	-3.661445	4.944	<0.001 ***
<b>Humboldtia brumonis</b>	incgr	4.4285582	3.2533069	5.3468710	3.592	<0.001 ***
	incgr0.33	-2.000926	-2.498270	-1.439176	2.877	<0.001 ***
	g15	-0.002749	-0.004724	-0.000946	2.615	<0.001 ***
	(Intercept)	-4.608481	-4.896508	-4.30573	9.009	<0.001 ***
<b>Myristica dactyloïdes</b>	dbh	0.0553601	0.0305054	0.0821755	7.983	0.006 **
	dbh2	-0.001083	-0.001554	-0.000642	8.470	0.010 **
	incgr	4.0284755	3.3842549	4.5965160	11.936	<0.001 ***
	incgr0.33	-3.449141	-3.779106	-2.945559	10.201	<0.001 ***
	(Intercept)	-4.089	-4.200	-3.987	5.637	<0.001 ***
<b>Knema attenuata</b>	incgr	2.595	1.692	3.303	3.002	<0.001 ***
	incgr0.33	-2.237	-2.680	-1.864	3.592	<0.001 ***
<b>Palaquium ellipticum</b>	Intercept	-5.41899	-5.73646	-5.18281	4.456	<0.001 ***
	dbh	0.05362	0.03938	0.06927	5.184	<0.001 ***
	incgr	-2.47407	-3.95682	-1.62175	2.967	<0.001 ***
	SlopeFactforte	0.27391	-0.01343	1.02762	2.369	0.068 .



## Résumé

La mortalité est un processus important de la dynamique des forêts, elle est pourtant moins étudié que la croissance. Ce mémoire présente un travail de modélisation de la mortalité des arbres en forêt tropicale humide. Les analyses sont réalisées à partir de données de suivi dendrométrique (recrutement, croissance et mortalité) d'un dispositif permanent d'échantillonnage situé dans les Ghâts occidentaux en Inde : le dispositif d'Uppangala. Plusieurs modèles logistiques ont été ajustés par des modèles linéaires généralisés et des modèles linéaires généralisés mixtes pour prendre en compte la répétition des mesures. Il ressort de cette étude que l'accroissement est le meilleur prédicteur de la mortalité pour ce dispositif et que la compétition comme le diamètre apportent une part explicative au modèle pour certaines espèces. La pente apparaît enfin comme un facteur augmentant la probabilité de mortalité des arbres. Les exemples de deux espèces émergentes ou de canopée : *Vateria indica* et *Palaquium ellipticum* montrent des mécanismes de mortalité liés à la pente à priori différents, plutôt régulière pour la première et plutôt irrégulière pour la seconde.

## Abstract

Tree mortality is a key process of forest dynamics. However, tree mortality studies are less frequent than growth studies. This study tackles the issue of modeling the tree mortality in a wet tropical forest located in the Indian Western Ghâts. Uppangala Permanent Sampling Plot has been monitored for 20 years and has been the subject of this study. Several logistic models have been assessed for the main 5 species and the 4 potential stratas of this stand. Generalized linear models and generalized linear mixed models have been used to calibrate the logistic models and take into account the repeated measurements correlations. These models show that past growth is the best mortality predictor on Uppangala's plots. Competition indexes and tree size are also used as predictor. Slope as also an influence on mortality in several species. The two emergent and canopy species *Vateria indica* and *Palaquium ellipticum* show different death patterns according to slope. An increase of the mortality is predicted with slope but it seems associated with regular mortality for *Vateria indica* and irregular mortality for *Palaquium ellipticum*.