



**HAL**  
open science

## **Cartographier le carbone stocké dans la végétation : perspectives pour la spatialisation d'un service écosystémique**

Solen Le Clec'H, Johan Oszwald, Nicolas Jegou, Simon Dufour, Pierre-André Cornillon, Luis Gonzaga, Izildinha Miranda, Michel Grimaldi, Valery Gond,  
Xavier Arnauld de Sartre

► **To cite this version:**

Solen Le Clec'H, Johan Oszwald, Nicolas Jegou, Simon Dufour, Pierre-André Cornillon, et al.. Cartographier le carbone stocké dans la végétation : perspectives pour la spatialisation d'un service écosystémique. Bois et Forêts des Tropiques, 2013, 67, pp.35-47. halshs-00923970

**HAL Id: halshs-00923970**

**<https://shs.hal.science/halshs-00923970v1>**

Submitted on 7 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Cartographier le carbone stocké dans la végétation : perspectives pour la spatialisation d'un Service Ecosystémique

LE CLEC'H Solen<sup>1</sup>, OSZWALD Johan<sup>1,2</sup>, JEGOU Nicolas<sup>3</sup>, DUFOUR Simon<sup>1</sup>, CORNILLON Pierre-André<sup>3</sup>, MIRANDA Izildinha<sup>4</sup>, GONZAGA Luiz<sup>4</sup>, GRIMALDI Michel<sup>5</sup>, GOND Valery<sup>6</sup> et ARNAULD DE SARTRE Xavier<sup>7</sup>.

1. Université Européenne de Bretagne - Rennes 2  
Laboratoire LETG Rennes – COSTEL (UMR CNRS 6554)  
5 Place Henri Le Moal, Rennes  
Contacts : solenleclech@hotmail.com ; [johan.oszwald@uhb.fr](mailto:johan.oszwald@uhb.fr) ; [simon.dufour@uhb.fr](mailto:simon.dufour@uhb.fr)

2. Laboratoire ECODIV (UPRES-EA 1293)  
Université de Rouen  
76821 Mont Saint Aignan

3. Université Européenne de Bretagne - Rennes 2  
Département de Mathématiques Appliquées au Sciences Sociales  
Contacts : [nicolas.jegou@uhb.fr](mailto:nicolas.jegou@uhb.fr)

4. Universidade Federal Rural da Amazonia (UFRA),  
2501 Av. Presidente Tancredo Neves, 66077-530 Belém (Pará), Brazil

5. Institut de Recherche pour le Développement (IRD)  
UMR 211 BIOEMCO  
Centre IRD France-Nord  
32, av. H. Varagnat, F-93143 Bondy cedex

6. CIRAD, upr Bsef, Biens et services des ecosystems forestiers tropicaux,  
Campus international de Baillarguet  
F-34398 Montpellier cedex 05  
Contact : [valery.gond@cirad.fr](mailto:valery.gond@cirad.fr)

7. Centre National de la Recherche Scientifique (CNRS), UMR 5603 SET  
Université de Pau et des Pays de l'Adour  
64000 PAU  
Contact : [xavier.arnauld@univ-pau.fr](mailto:xavier.arnauld@univ-pau.fr)

## Résumé

Le Millenium Ecosystem Assessment (MEA), puis les REDD (Reducing Emissions from Deforestation and forest Degradation) et les REDD+, préconisent le développement des approches permettant de quantifier et de spatialiser les services écosystémiques (SE) afin de mettre en œuvre des pratiques et des politiques de gestion environnementale plus adaptées. La cartographie des SE apparaît ainsi comme un outil majeur des espaces à forts enjeux environnementaux. Cependant, elle souffre encore de certaines limitations.

C'est le cas du stock de carbone dans la biomasse végétale. À l'échelle d'une localité d'Amazonie brésilienne de 175 km<sup>2</sup>, nous avons cartographié, avec une résolution spatiale de 30x30m, cette fonction écologique. Afin de quantifier ces stocks, des mesures de biomasse arborée et arbustive au sein de 45 « points » et des données géographiques obtenues par télédétection sont mises en relation. Pour cela, deux méthodes statistiques sont testées : l'arbre de décision et la régression linéaire multiple. Les résultats statistiques de chacune de ces méthodes sont présentés, permettant d'en comprendre les intérêts et les inconvénients. La qualité d'ajustement de ces modèles est testée. Si l'arbre de décision décrit mieux le rôle des variables explicatives, la régression linéaire multiple permet une prédiction beaucoup plus efficace. Elle rend alors davantage compte de la variabilité spatiale au sein de chaque type d'occupation du sol. Cette méthode fait apparaître à l'échelle de la ferme des phénomènes spécifiques au territoire étudié. Cela permet de retranscrire simplement le résultat d'un processus écologique tout en le mettant en relation avec les activités anthropiques. Cette étude permet donc d'illustrer l'importance des choix méthodologiques afin d'obtenir la cartographie d'un processus.

## Introduction

L'Amazonie brésilienne est une région où les enjeux en termes de biodiversité sont bien connus (PAVE, 2010), ce qui en a fait un emblème des politiques de conservation de la biodiversité depuis les années 1980. Les espaces forestiers amazoniens rendent également un grand nombre de services aux sociétés : ils fournissent du bois de chauffage et de construction, jouent un rôle dans les cycles de l'eau et du carbone... Or l'Amazonie brésilienne est devenue, depuis les années 1970, une des régions du globe à connaître des taux de déforestation parmi les plus élevés (LENA, 1999 ; EVANS *et al.*, 2001; FEARNSSIDE, 2005). Les activités de déboisement ont pour but la mise en valeur des nouvelles terres par l'élevage et l'agriculture. Les conséquences sur le milieu ne sont donc pas négligeables, que ce soit sur les écosystèmes, leurs fonctionnements ou les processus associés.

Dans un contexte de déforestation, un des services écosystémiques (SE) les plus mis en avant par les organismes internationaux est la régulation du climat via la capacité des écosystèmes à séquestrer le carbone. Le service de séquestration du carbone présente de forts enjeux puisque la végétation des écosystèmes terrestres stocke plus de carbone que l'atmosphère, et ce particulièrement dans les forêts tropicales (SILVER *et al.*, 2000). Les végétaux absorbent ainsi une

partie du dioxyde de carbone contenu dans l'atmosphère pour assurer leur activité photosynthétique. Ce processus de séquestration est de mieux en mieux appréhendé, comme en témoigne la multiplication des dispositifs internationaux visant à protéger les espaces considérés comme des puits de carbone, notamment les régions forestières afin que celles-ci n'en deviennent pas des sources. De fait, la question du stock de carbone dans la végétation est souvent traitée dans les études scientifiques (BROWN, 1997). Les représentations cartographiques associées, bien que plus rares, se multiplient également, d'autant que des dispositifs internationaux tels que les REDD (Reducing Emission from Deforestation and forest Degradation) en préconisent le développement.

La plupart des approches utilisent l'occupation du sol : typologie de l'occupation des sols spatialisée par télédétection (PONCE-HERNANDEZ, 2007 ; HAPSARI, 2010) et inventaires forestiers (FUKUDA *et al.*, 2012). Les auteurs appliquent des équations allométriques fondées sur la mesure des arbres, notamment de la fermeture de la canopée (NELSON *et al.*, 1998 ; WAUTERS *et al.*, 2007). La prédominance de l'occupation du sol comme donnée d'entrée s'explique par plusieurs facteurs. En effet, l'occupation du sol semble être un bon indicateur des variations dans le stock de carbone de la biomasse végétale puisque celui-ci est sensible à l'occupation et la mise en valeur des sols. De plus, cette donnée est facile à acquérir par les outils satellitaires. De fait, le Millenium Ecosystem Assessment (MEA), en soulignant l'importance de cet indicateur dans l'estimation de la dégradation des SE, préconise son utilisation. Malgré tout, il peut être envisagé que d'autres facteurs puissent intervenir dans les variations du taux de carbone stocké. Un jeu de données plus complet, à la fois plus diversifié en termes de variables explicatives et plus précis en termes d'occupation du sol permettrait de mieux expliquer et quantifier le phénomène étudié, sans le simplifier.

L'objectif de cet article est donc de cartographier le stock de carbone dans la biomasse arborée et arbustive à l'échelle locale (175 km<sup>2</sup>) avec une précision spatiale de 30x30m, niveau de précision rarement atteint dans la littérature. Cet objectif nécessite de résoudre deux questions méthodologiques importantes : celle des données d'entrée et celle des modèles statistiques à utiliser. Deux méthodes sont testées afin de modéliser les stocks de carbone mesurés sur le terrain : la régression linéaire et l'arbre de décision. Ces deux modèles ont été retenus notamment du fait de leur capacité à prédire des valeurs quantitatives (stock de carbone) à partir de données d'entrées quantitatives et qualitatives (CORNILLON *et al.*, 2010).

Il est important de noter que cette étude vise à cartographier le stock de carbone et non son stockage. La différence est importante. En effet, le stock décrit un état, résultant de la photosynthèse et non un processus dynamique réalisant le service écosystémique lui-même. Si certaines études choisissent de quantifier les stocks de carbone en valeurs économiques, la cartographie réalisée ici se fonde sur une analyse biophysique avec le recours à des unités ayant écologiquement un sens (NAIDOO *et al.*, 2006 ; RUTLEDGE *et al.*, 2010). Il s'agit donc de la quantité de carbone stocké dans la biomasse arborée et arbustive à un instant *t*. Ce stock est susceptible de connaître d'importantes variations dans l'espace et dans le temps, en fonction des

usages, permettant ainsi de préciser l'importance du déstockage (à partir de la forêt primaire) ou du stockage (par exemple, depuis les pâturages jusqu'aux forêts secondaires). Les cartographies du stock de carbone peuvent trouver une application dans le contexte de dispositifs tels que les REDD et REDD+ dans la mesure où elles fournissent un état des lieux de la quantité de carbone présent dans la végétation. Cela en fait un outil indispensable pour développer une approche critique des nouveaux outils de gestion des territoires comme les paiements pour des services écosystémiques.

### **Zone d'étude**

Le site de Pacajá, sélectionné pour l'étude, appartient à la région de la Transamazonienne (Etat du Pará en Amazonie brésilienne - figure 1A et 1B). Celle-ci constitue une des zones forestières où la déforestation est particulièrement rapide ces vingt dernières années (ARNAULD DE SARTRE et BERODULAY, 2011). Le front de déforestation s'étend à partir de la route fédérale BR 230 (aussi appelée Transamazonienne) et se développe le long de pistes secondaires dites en « arrêtes de poisson » (c'est-à-dire qui, tous les cinq kilomètres, se déploient de part et d'autre de la route principale - figure 1C et 1D). Le site étudié appartient au *município* de Pacajá et se situe le long d'une de ces pistes secondaires (la piste 338 Nord). Ce site connaît, comme la plus grande partie des marges orientales et méridionales de l'Amazonie, un taux de déforestation très élevé avec l'installation de colons agricoles qui pratiquent essentiellement l'élevage bovin.

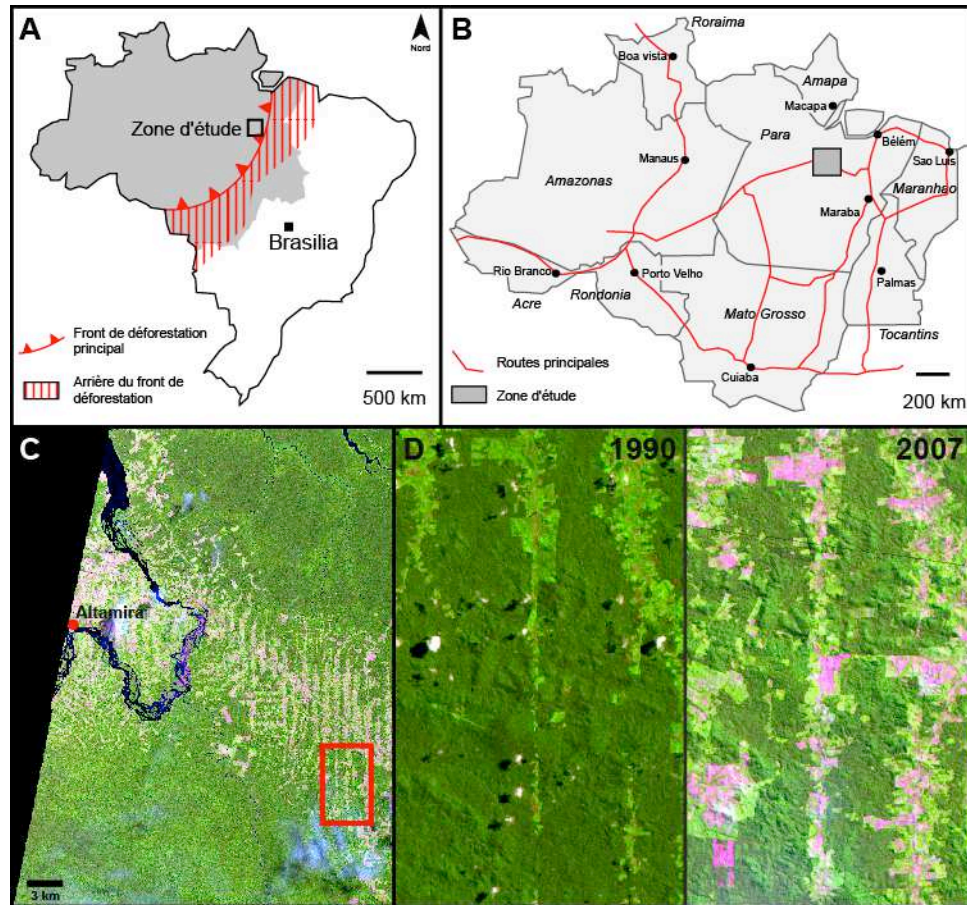


Figure 1 : A. Carte du Brésil – en gris apparaissent les régions amazoniennes. B. Carte de l’Amazonie brésilienne légale avec ses différents États – en gris plus sombre apparaît la région de Pacajá. C. Carte de la région Pacajá (image Landsat Thematic Mapper (TM) fausse couleur de 2007, rouge / Proche InfraRouge (PIR) / Moyen InfraRouge(MIR)) et localisation de la région de Pacajá. D. Images Landsat TM de 1990 et de 2007 de la zone d’étude (rouge/PIR/MIR).

## Données

Deux jeux de données différents ont été utilisés pour construire le modèle de prévision du stock de carbone dans la biomasse arborée et arbustive : les données relevées *in situ*, qui constituent la variable à expliquer, et les données obtenues par télédétection qui sont des variables explicatives (Tableau 1). Un modèle prédictif est construit à partir des données observées et a été utilisé pour inférer les valeurs de carbone stocké sur la zone de Pacajá à partir des seules données de télédétection. Les variations locales sont potentiellement importantes et l’information se doit alors d’être la plus détaillée possible (BURKHARD, 2009). Du fait de l’utilisation d’images satellites Landsat TM et du MNT Aster (résolution 30x30m), l’unité de base de la carte correspond à un pixel de 900 m<sup>2</sup>.

Type de variable	Nom	Sources
<b>Variables à expliquer</b>	Stock de carbone dans la végétation arborée et arbustive (Mg/ha)	Relevés terrain
<b>Variables explicatives</b>	<b>Occupation des sols :</b> - Historique (1986 - 2007) - Actuelle	<b>Images Landsat TM (30x30 m.)</b>  Traitement à partir des bandes Rouge, Proche InfraRouge, Moyen InfraRouge
	<b>MNT et ses dérivés :</b> - Altitude (m) - Pente (%) - Topographie - Distance à l'eau	<b>MNT Aster (30x30 m)</b>  Traitement sous ArcGis 9.2 et 9.3 avec les extensions Spatial Analyst et Corridor Designer

Tableau 1 : récapitulatif des variables utilisées dans les analyses statistiques.

Les données *in situ* proviennent de campagnes de terrain effectuées en 2007-2008 sur le site de Pacajá. L'échantillon comprend neuf fermes sélectionnées de manière à représenter les différents types d'exploitations présentes sur le site (Figure 2). Pour chacune des fermes échantillonnées, un transect a été défini. Sur celui-ci, cinq points d'échantillonnages ont été positionnés à intervalle régulier, à environ 200 m les uns des autres (OSZWALD *et al.*, 2011). Suite à une analyse statistique préliminaire, seuls 43 points sont utilisés, 2 d'entre eux présentant des valeurs aberrantes.





Figure 2 : Occupation du sol du site de Pacajá et répartition spatiale des 45 points échantillonnés sur le terrain.

Nous disposons ainsi pour 43 points de données sur la nature et la composition chimique des sols, sur la végétation, sa biomasse et le stock de carbone. En chacun de ces points, la biomasse épigée des arbres (diamètre à hauteur de poitrine (dbh)  $\geq 10$  cm) et des arbustes (dbh  $< 10$  cm et hauteur  $> 2,0$  m) a été évaluée par l'intermédiaire d'équations allométriques disponibles pour les forêts (HIGUCHI *et al.*, 1998 ; GERWING, 2002) et jachères (NELSON *et al.*, 1999), après mesure du diamètre et estimation de la hauteur de chaque plante sur des parcelles de  $50 \times 10$  m<sup>2</sup> et  $50 \times 5$  m<sup>2</sup>, respectivement, et de la mesure de la teneur en eau à 70°C d'une aliquote de biomasse correspondante (toutes les plantes de dbh  $\leq 5$  cm), ou (dbh  $\geq 5$  cm) en appliquant le facteur 0,603 en accord avec HIGUCHI *et al.* (1998). En l'absence d'analyse de teneur en carbone sur la végétation, le facteur 0,5 a été utilisé pour estimer la masse de carbone à partir de la biomasse végétale sèche (MARKEWITZ *et al.*, 2004). La masse de carbone est exprimée en mégagramme par hectare (Mg/ha) ou tonne par hectare.



Les réserves de carbone stocké dans la biomasse arborée et arbustive étant étroitement liées aux conditions environnementales, les variables explicatives choisies ont été construites à partir de données obtenues par télédétection renseignant donc l'occupation du sol (dynamique spatio-temporelle) et la topographie. Elles sont issues du traitement des images Landsat TM de 1986, 1996, 2001 et 2007 effectué sous ENVI ainsi que du traitement du Modèle Numérique de Terrain (MNT) Aster réalisé sous ArcGis 9.2. Ces données sont connues pour l'ensemble du site de Pacajá (OSZWALD *et al.*, 2012). L'ensemble des données spatialisées obtenues par télédétection ont été extraites pour les 43 individus renseignés *in situ*. Les caractéristiques topographiques dérivées du MNT Aster (altitude, pente, topographie et distance à l'eau) et l'occupation des sols pour l'année la plus récente, mais aussi de manière rétrospective (1986-2007).

Les données de topographie renseignent en tout point la zone d'étude l'altitude (en mètre). Les altitudes de la zone de Pacajá s'échelonnent de 13 à 359 mètres. Les pentes synthétisent la différence d'altitude entre deux pixels adjacents et sont fournies en pourcentage. Ces deux variables sont quantitatives et sont traitées comme des données brutes continues. La variable « topographie » correspond à une caractérisation synthétique du contexte topographique comprenant quatre modalités : fonds de vallées, sommets de collines, zones de fortes pentes et zones de faibles pentes. Enfin, le MNT a permis de déduire une cartographie du réseau hydrographique et ensuite de déterminer une distance aux cours d'eau (0 à 100 mètres, 100 à 200 mètres, 200 à 300 mètres, 300 à 500 mètres et plus de 500 mètres). L'envergure des cours d'eau n'est pas prise en compte puisqu'il n'y a pas de différence de taille entre les cours d'eau de la zone d'étude.

Par ailleurs, l'analyse et le traitement des images satellites Landsat TM a également permis de déterminer six types d'occupation des sols pour 2007 (figure 3). De plus, la cartographie des types d'occupation des sols a également été étendue à l'ensemble des images satellites du jeu de données (de 1986 à 2007).

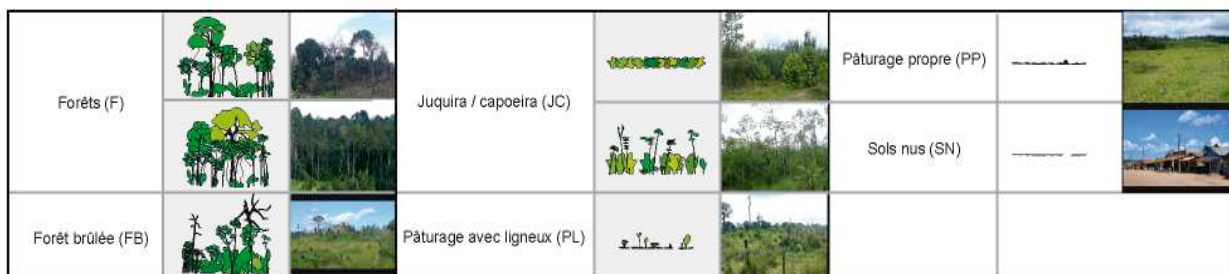


Figure 3 : Typologie des éléments paysagers caractérisée à partir de l'analyse des images satellites entre 1990 et 2007. Deux types de forêt ont été représentées, la forêt exploitée (photo du haut) et forêt dense (image du bas). La capoeira (image du haut) est une jachère dominée par les herbacées alors que dans la juquira (image du bas) la strate arbustive puis arborée dominant.

À partir des cartes d'occupation des sols obtenues pour les quatre dates, cinq classes de trajectoires d'occupation des sols ont été déterminées (OSZWALD *et al.*, 2012). La classe 1 représente une structure homogène et forestière. La classe 2 correspond à des couvertures du sol avec quelques évolutions agricoles fortement polarisées par la proximité au réseau de pistes. Une variable de distance au réseau de piste n'est pas retenue du fait de leur forte concordance à la dynamique de l'occupation des sols. La classe 3 présente une rupture par rapport aux deux classes précédentes. En effet, l'emprise agricole sur la ferme est plus précoce et les dynamiques agricoles sont beaucoup plus intenses à partir des années 2000. La classe 4 correspond à des fermes exploitées antérieurement aux années 1990 avec une forte emprise agricole le long du réseau de pistes. Enfin, la classe 5 correspond à une dynamique agricole d'élevage extensif. Cette donnée paraît particulièrement intéressante pour déterminer l'impact des dynamiques paysagères sur le stock de carbone relevé en 2007, particulièrement dans un contexte de front de déforestation actif.

## **Méthodes**

D'un point de vue statistique, la démarche consiste à construire un modèle en croisant des données de télédétection et des observations terrain, le but étant ensuite d'utiliser ce modèle pour prédire les quantités de carbone stocké à partir des seules données de télédétection. Les valeurs prédites sont alors utilisées sous SIG afin d'obtenir une cartographie de la zone. Cette modélisation a deux objectifs. D'une part, elle vise à expliquer les variations du stock de carbone à partir des données sélectionnées. Construire le modèle consiste en effet à sélectionner de manière objective certaines variables, à estimer ensuite les paramètres, afin de comprendre les facteurs influençant les processus écologiques mis en jeu. D'autre part, en prédisant des valeurs sur l'ensemble de la localité, le modèle permet d'obtenir une cartographie pour l'ensemble d'un secteur à partir de quelques points d'échantillonnage. Pour modéliser une variable quantitative à partir de variables qualitatives et quantitatives les méthodes les plus répandues sont les arbres de régression (Algorithme Classification And Regression Trees – CART) et la régression linéaire multiple.

Les deux ont été mises en œuvre et comparées par validation croisée. Le principe de la validation croisée consiste à séparer l'ensemble des données en un échantillon d'apprentissage sur lequel on construit chaque modèle et d'un échantillon test sur lequel le modèle est appliqué pour prédire les quantités de carbone stocké. Une erreur de prévision mesurant l'écart entre les valeurs prédites sur l'échantillon test et les valeurs observées sur le terrain est calculée pour chaque modèle. Dans la procédure mise en place, 39 individus sur 43 ont été tirés aléatoirement pour former l'échantillon d'apprentissage et au total, 100 relances ont été faites permettant ainsi de comparer les vertus prédictives des deux modèles via la confrontation des erreurs de prévision cumulées.

### ***Arbre de décision (algorithme Classification And Regression Trees - CART)***

Le principe d'un modèle par arbre de décision est de partitionner l'espace des variables explicatives en différentes classes qui ne se recouvrent pas ; classes au sein desquelles la variabilité entre individus est faible du point de vue de la variable à expliquer (Figure 4). Ainsi, des groupes d'individus homogènes pour la variable « stock de carbone dans la biomasse arborée et arbustive » sont formés, la valeur ajustée correspondant à la moyenne des observations pour les individus du groupe. La méthode standard consiste en l'application d'un algorithme séquentiel fondé sur la méthode CART (BREIMAN *et al.*, 1984). Cet algorithme stoppe son partitionnement lorsque le fait de rajouter des coupures n'apporte plus une information significative supplémentaire. A l'issue de la procédure, le partitionnement se fait uniquement avec certaines variables. En ce sens, la procédure fait aussi office de sélection de variables. L'un de ses avantages est qu'elle propose une représentation du modèle retenu sous la forme d'un arbre facile à interpréter.

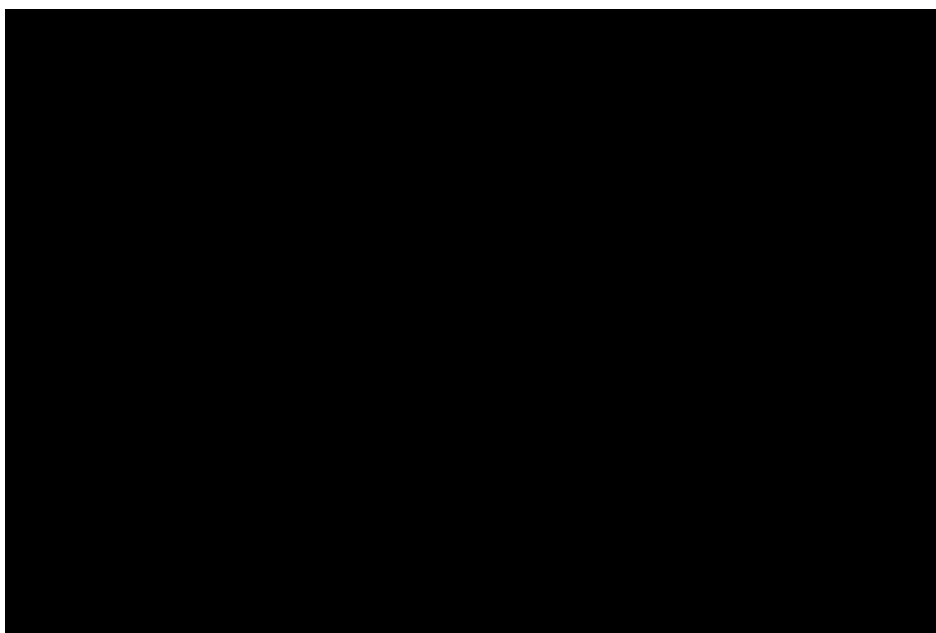


Figure 4 : Représentation schématique du principe de l'arbre de décision avec  $X_1$  et  $X_2$ , des variables explicatives et  $Y$  la variable à expliquer (les différentes couleurs représentant les différentes classes issues du modèle).

### ***Régression linéaire multiple***

La régression linéaire multiple est une généralisation de la régression linéaire simple. Le modèle est construit en combinant linéairement les variables explicatives. Les paramètres estimés par moindres carrés correspondent aux coefficients associés aux variables dans cette combinaison (CORNILLON, 2011). Les paramètres estimés par le modèle confèrent ainsi un poids à chacune des variables selon leur influence. L'estimation du modèle complet montre que plusieurs paramètres sont significativement différents de 0 (Annexe 1). Cela révèle que certaines variables n'influencent pas de manière significative la variable stock de carbone. Une procédure de choix de variables a ainsi été mise en place. Le principe de cette procédure consiste à satisfaire un critère qui réalise un compromis entre la qualité de l'ajustement (mesurée par la variance

résiduelle) et un nombre raisonnable de variables. Formellement, le critère s'écrit sous la forme générale :

$$\text{Variance résiduelle} + \varnothing (\text{nombre de paramètres})$$

La variance résiduelle (écart quadratique moyen entre valeurs prédites et valeurs observées) diminue mécaniquement lorsque l'on ajoute des variables au modèle : chercher à la minimiser conduira toujours à retenir le modèle complet. Le second terme au contraire est une fonction croissante du nombre de variables. Le modèle minimisant un critère de ce type réalise donc un équilibre entre fidélité aux données et pénalisation d'un trop grand nombre de variables. Plusieurs formulations existent pour la fonction  $\varnothing$  définissant ainsi les critères associés. Les plus classiques sont le critère AIC (Akaike's Information Criterion, AKAIKE, 1973), le critère BIC (Bayesian Information Criterion (SCHWARZ, 1978) ou le critère Cp (MALLOWS, 1973). Le critère AIC est d'ordinaire utilisé en régression non paramétrique, aussi ce sont les deux autres critères qui ont été utilisés et ceux-ci retiennent le même modèle (LIU, 2001 ; Annexe 2).

Une fois les modèles réalisés, des valeurs sont prédites pour l'ensemble de la zone de Pacajá, et sont injectées sous SIG avec ArcGis 9.3 afin d'obtenir la cartographie des quantités de carbone stocké dans la végétation.

## Résultats

L'estimation de la quantité de carbone stocké dans la biomasse arborée et arbustive et sa distribution statistique sont très proches quel que soit le modèle choisi (Tableau 2).

Modèle statistique	Quantité totale (Mg)	Quantité moyenne (Mg)	Quantité minimale (Mg/ha)	Quantité maximale (Mg/ha)
Arbre de décision	46 657 360	234.5	0.31	295.0
Régression linéaire	42 405 547	213.1	0.0	381.1

Tableau 2 : Valeurs de carbone stocké dans la végétation arborée et arbustive obtenues par les deux modèles.

Les valeurs déterminées par les deux modèles sont en moyenne plus élevées que celles obtenues par d'autres auteurs concernant l'Amazonie brésilienne (sur d'autres sites ou à des échelles spatiales différentes - Tableau 3).

Sources	Echelle/site	Quantité moyenne estimée
Achard <i>et al.</i> (2004)	Amazonie brésilienne	186 t / ha
Brown (1997)	Amazonie brésilienne	186 t / ha

Rice <i>et al.</i> (2004)	Forêt Nationale de Tapajo, Pará, Brésil	143.7 Mg / ha
Salimon <i>et al.</i> (2010)	Acre, Brésil	246 Mg / ha

Tableau 3 : Estimations des quantités de carbone stocké dans la végétation arborée et/ou arbustive dans d'autres études menées sur l'Amazonie brésilienne.

En effet, le stock de carbone mesuré ici n'est pas uniquement celui de la biomasse arborée, comme dans la plupart des études qui ont déjà été réalisées, mais concerne également la biomasse arbustive, ce qui tend à mieux appréhender le carbone stocké hors du sol. De plus, le jeu de donnée utilisé dans cet article a permis une estimation plus fiable de la quantité de carbone stocké car il ne se fonde pas uniquement sur l'occupation du sol, notamment par la seule opposition "forêt / non forêt". Enfin, la plupart des autres études sont menées à une échelle plus large, souvent régionale, ce qui pose un problème de cohérence dans les échelles utilisées, puisque les processus physiques permettant d'expliquer le carbone stocké ont cours à des échelles locales, voire à celle de l'arbre. Ainsi, plus le décalage d'échelle entre les données utilisées est important, plus les risques d'induire des erreurs d'estimation sont élevés.

### Arbre de décision

Le modèle par arbre de décision est simple puisqu'il ne retient qu'une seule variable, l'« occupation du sol » (Figure 5). Les autres variables présentées préalablement n'interviennent pas dans la prédiction du stock de carbone pour cette zone.

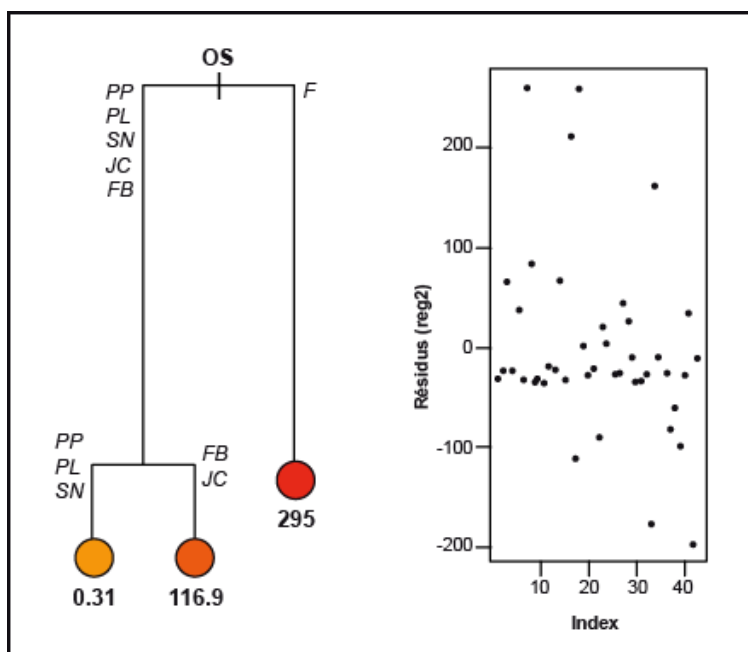


Figure 5 : Arbre de décision du modèle expliquant la quantité de carbone stocké dans la biomasse arborée et arbustive pour les 43 points échantillonnés (mise en œuvre avec le logiciel R,

package rpart (CORNILLON *et al.*, 2010)) et résidus du modèle (OS : Occupation du Sol, pour les autres acronymes cf. figure 2).

La qualité d'ajustement du modèle mesurée par le  $R^2$  est de 51 %. Ce coefficient est significativement différent de zéro et sa p-value vaut 4.313e-06. Le modèle distingue d'abord les pixels de forêts des autres types d'occupation des sols. Une seconde division apparaît dans cette dernière catégorie où les valeurs de carbone sont significativement différentes pour les sites arborés (forêts brûlées et jachères et/ou friches) et les sites agricoles (pâturages et sols nus). Cette distinction paraît évidente du fait que les forêts brûlées ont déstocké une grande partie du carbone anciennement stocké sous l'action des brûlis successifs accomplis par les agriculteurs afin de mettre en place les futures exploitations agricoles (Photo 1).



Photo 1 : Brûlis de la couverture forestière accomplis par des agriculteurs. Entre les troncs calcinés sont visibles des plants de riz pluvial (photo J. Oszwald)

Cependant, il est très intéressant de noter que ce facteur est associé aux espaces de jachères et de friches. Les pixels de pâturages et de sols nus correspondent à des zones fortement anthropisées et stockent environ 1000 fois moins de carbone que les pixels de forêt dense. Les quantités de carbone stocké dans les pixels correspondant à des zones liées à une dégradation forestière moins avancée (forêt brûlée et friches) représentent environ le tiers des quantités stockées par rapport aux forêts denses.

### ***Régression linéaire multiple***

La procédure de sélection de variables, implémentée sous R, conduit à garder le modèle retenant deux variables : « Occupation du Sol » et « Distance à l'eau (plus de 300 mètres) » (Annexe 2). La qualité d'ajustement du modèle mesurée par le  $R^2$  est de 60 %. L'ajustement est donc meilleur que pour l'arbre de décision (Tableau 4).

<b>Modèle statistique</b>	<b>R<sup>2</sup></b>	<b>Validation croisée</b>
Arbre de décision	51%	66 410
Régression linéaire multiple	60%	37 768

Tableau 4 : Comparaison statistique des deux méthodes.

Une procédure de validation croisée a été exécutée. L'erreur de prévision cumulée est 46% supérieure dans le cas de l'arbre de décision (Tableau 4). Le modèle de régression semble donc plus performant pour prévoir la variable « Stock de carbone ». Malgré tout, l'arbre de décision présente l'intérêt d'être facile à comprendre et à interpréter avec une hiérarchisation visuelle des variables, ce qui en fait un outil pédagogique très intéressant et facile à s'approprier. Ces analyses ont permis une première identification des données nécessaires à la construction d'un modèle explicatif et prédictif d'un processus environnemental à l'origine d'un SE de régulation climatique.

### ***Résultats cartographiques***

Pour chacun des deux modèles proposés, la cartographie de la localité de Pacajá ainsi que celle de trois fermes sélectionnées de par leur situation géographique (une au nord, une au centre et une au sud) sont présentées (Figure 6).



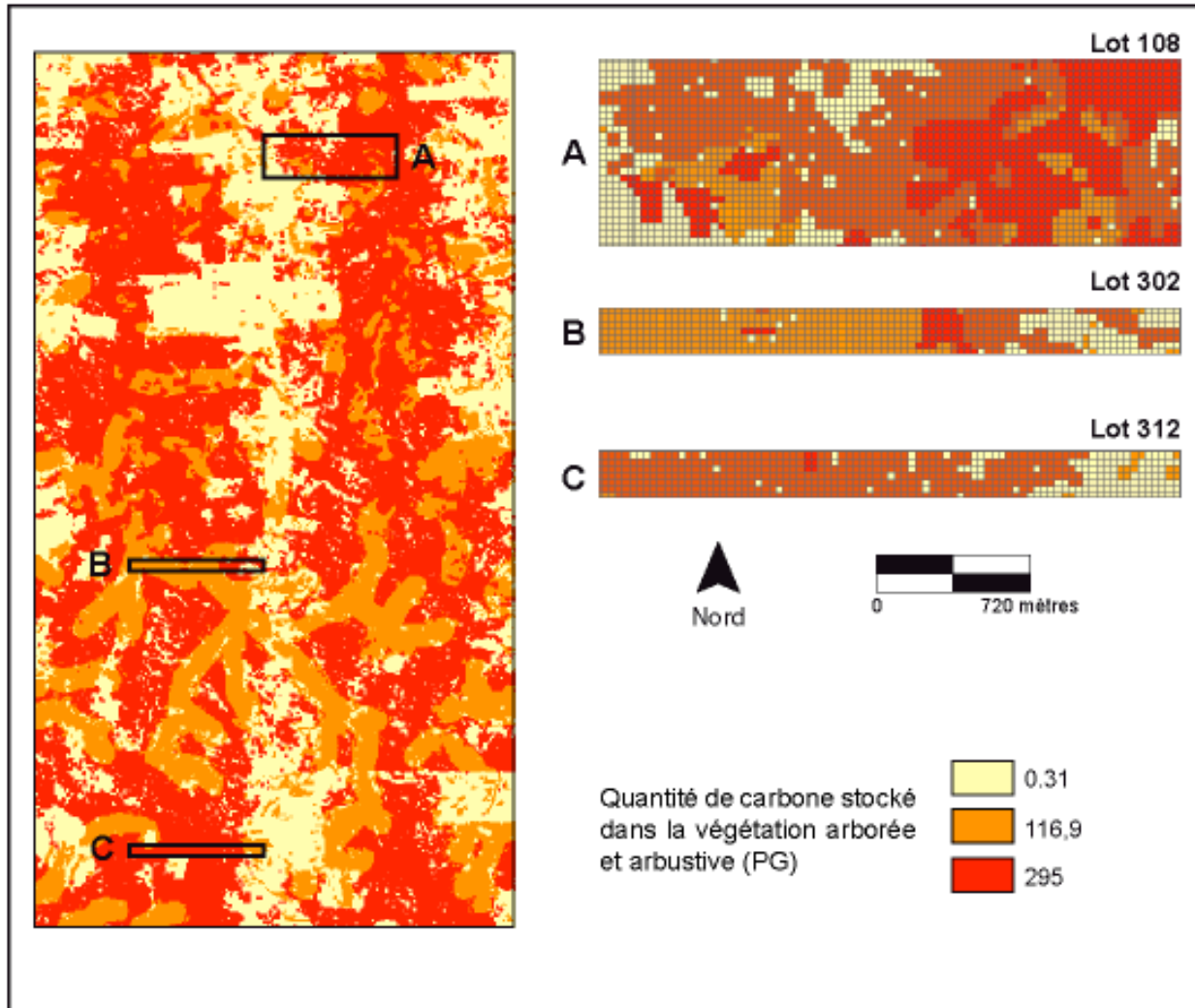


Figure 6 : Carte du carbone stocké dans la végétation arbustive et arborée dans la zone de Pacajá (zoom sur trois fermes à titre d'exemple) inférées à partir du modèle d'arbre de décision.

Les stocks de carbone les plus élevés se situent dans les zones les plus boisées (forêts denses et forêts brûlées) alors que les quantités les plus faibles, comprises entre 14,52 et 87,51 Mg/ha, correspondent aux fronts de déforestation de ces dernières années (Figure 6). Ce résultat est fortement corrélé à l'arbre de décision qui prévoyait les plus fortes valeurs de stock de carbone dans les espaces forestiers. À l'échelle des fermes, de fortes disparités de quantités de carbone stocké peuvent être observées. De manière sensible, les quantités de carbone stocké augmentent depuis le front principal de déforestation, où se situe le réseau de pistes, vers la marge opposée de la ferme (Figure 7).

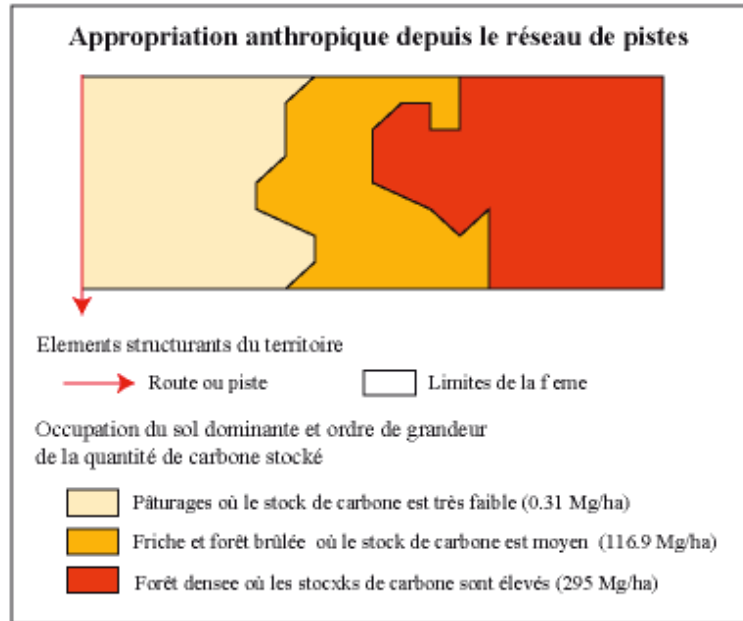


Figure 7 : Dynamique des structures et d'occupation des sols observées à l'échelle des fermes en relation avec les valeurs de carbone stocké dans la biomasse arborée et arbustive

Une ferme peut se scinder en trois espaces limitrophes ayant des caractéristiques de l'emprise anthropique différentes, ce qui induit un impact non négligeable sur les mesures de carbone stocké. En effet, les espaces situés en bordure du réseau de pistes ont été déboisés plus anciennement et correspondent à des espaces cultivés depuis les années 1990. Dans ces zones majoritairement pâturées, la végétation arborée et arbustive est peu présente. Les quantités de carbone stocké sont donc plus faibles. En s'éloignant de la piste, l'emprise anthropique est moins marquée, mais reste présente. Cela amène à la constitution d'une mosaïque d'espaces pâturés avec des patches de *juquira* et de *capoeira*. Ce sont des espaces de transition marqués par une forte hétérogénéité de la répartition spatiale des quantités de carbone stocké. L'emprise anthropique étant faible sur les zones les plus éloignées de la route, les couverts végétaux sont majoritairement forestiers et homogènes. Les quantités de carbone stocké y sont, par conséquent, les plus importantes.

La carte produite à partir du modèle de régression linéaire présente des valeurs de stockage de carbone plus différenciées sur l'ensemble de la zone de Pacajá que celles prédites par la méthode d'arbre de décision (Figure 8). Les valeurs prédites sont plus nombreuses, permettant ainsi de représenter plus finement les variations de quantité de carbone stocké.

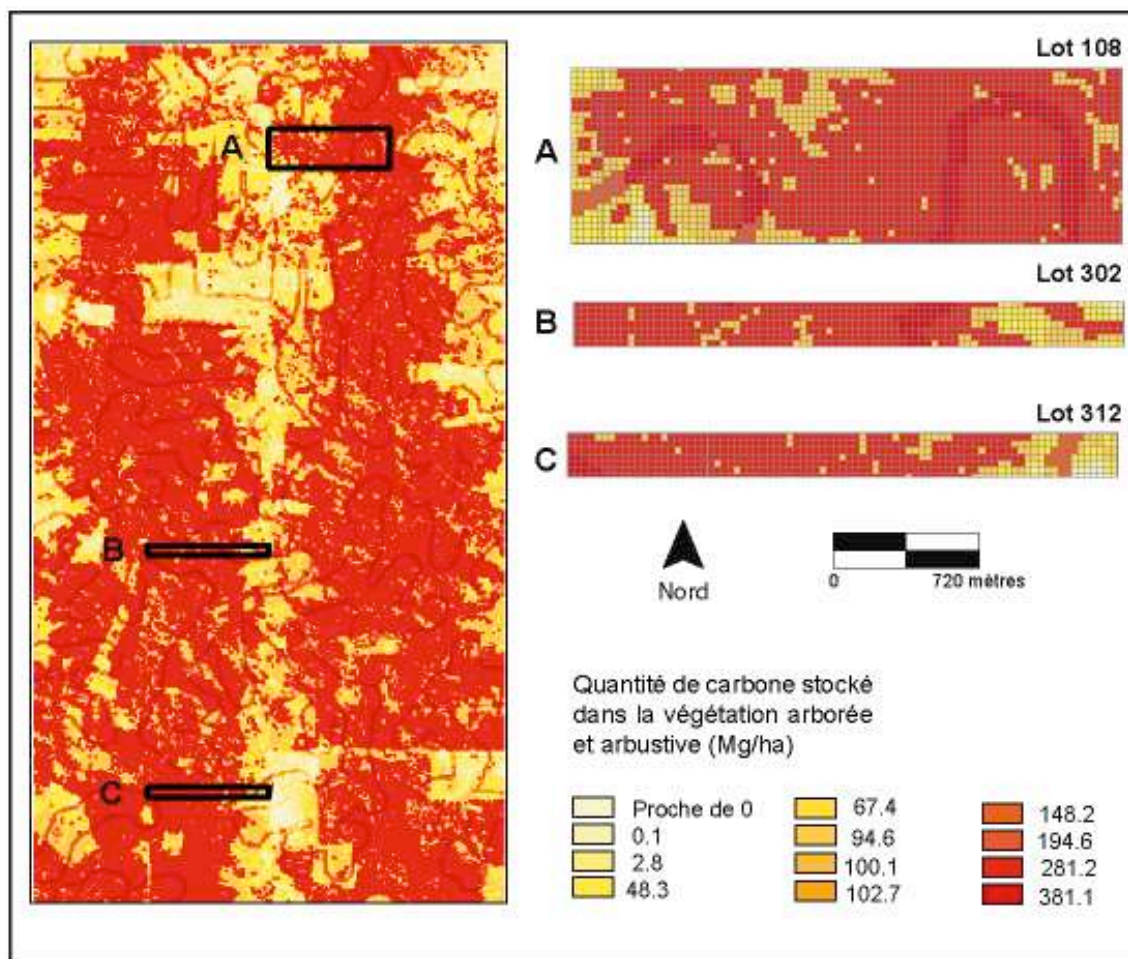


Figure 8 : Carte du carbone stocké dans la végétation arbustive et arborée dans la zone de Pacajá (zoom sur trois zones à titre d'exemple) inférées à partir du modèle de régression linéaire multiple.

A l'échelle de la localité, la carte permet de distinguer de fortes quantités de carbone dans les zones les plus boisées (forêts denses et forêts brûlées) et des quantités beaucoup plus faibles, au niveau du front de déforestation. La cartographie laisse apparaître un nouveau phénomène que la figure 6 ne permet pas de visualiser : le réseau hydrographique est mis en évidence avec des quantités de carbone stocké plus importantes. Ce phénomène s'explique par la présence de forêts galeries. À l'échelle des fermes, les forêts galeries ressortent nettement avec des quantités de carbone atteignant 381,1 Mg/ha. Les forêts galeries sont d'autant plus présentes que le Code Forestier brésilien préconise aux agriculteurs de planter ou de replanter des arbres aux abords des cours d'eau. Cependant, les quantités de carbone stocké dans ces forêts galeries varient fortement aux échelles locales, comme nous pouvons le remarquer dans la ferme A (Figure 8). Les forêts galeries situées à l'est de la ferme stockent des quantités de carbone beaucoup plus fortes que celles situées dans les espaces pâturés de l'ouest de la ferme.

En effet, les forêts galeries se développant dans un contexte pâturé sont d'avantage soumises à des brûlis successifs dus à l'entretien du pâturage visant à limiter le développement de la végétation aux abords de la forêt galerie. De plus, les troupeaux se trouvant dans ces pâturages impactent également ces forêts galeries et entraînent leur dégradation (structurelles et qualitatives). De plus, la cartographie issue de la régression linéaire multiple permet de mieux appréhender la diversité du stock de carbone dans les espaces pâturés. Ces variations sont dues à une forte hétérogénéité des types de pâturage avec la présence de pâturage propres, de pâturages avec ligneux et de sols nus (Figure 8A).

## **Discussion et perspectives**

L'approche étant fondée sur une prévision statistique, il existe des incertitudes. Il peut être envisagé d'intégrer d'autres données afin d'améliorer la fiabilité des modèles statistiques (indices radiométriques tels que le NDVI, réseaux d'infrastructures tels que les routes). Cependant, au regard de la bibliographie récente sur le sujet, les modèles proposés offrent des résultats satisfaisants et l'ajout de nouvelles variables n'apportait pas d'amélioration des modèles (LECOINTE, 2006 ; ASNER, *et al.*, 2012). Plus encore, l'augmentation du nombre de variables, notamment si elles sont corrélées spatialement, risque d'entraîner une augmentation mécanique de la qualité des prévisions du modèle sans que celle-ci soit statistiquement significative. Les deux méthodes statistiques appliquées démontrent l'importance des choix méthodologiques puisque les quantités de carbone prédites dépendent du modèle utilisé. La cartographie issue de l'arbre de décision permet de mettre en relation les quantités de carbone stocké dans la végétation arborée et arbustive et les grandes tendances historiques d'appropriation de la terre. La cartographie réalisée par la régression linéaire permet une analyse beaucoup plus fine permettant de rendre d'avantage compte de la variabilité spatiale au sein de chaque type d'usage du sol. Cela est dû à la capacité prédictive supérieure de cette méthode et à la présence d'une seconde variable explicative qui permet d'affiner l'analyse. Enfin, sur ces espaces récemment déforestés, il apparaît que la variable présentant les dynamiques historiques de l'occupation des sols est peu pertinente. En effet, le front de déforestation ayant moins de 15 ans à Pacajá, la cartographie actuelle de l'occupation des sols transcrit pleinement les dynamiques en cours. De plus, l'apport d'une typologie détaillée permettant de rendre compte le plus finement possible des différents types d'occupation des sols semble indispensable. Ainsi, l'utilisation d'images satellites Landsat TM est pertinente, que ce soit en termes d'identification des types d'occupation des sols, mais aussi en termes de résolution spatiale. En effet, en travaillant à partir d'images à haute résolution spatiale, les cartographies produites permettent d'identifier des variations du stock de carbone aux échelles fines et ainsi de mieux appréhender la relation statistique entretenue entre les relevés terrain, donc les processus physiques, et les données télédéteectées. Ainsi, les échelles locales constituent une échelle d'analyse importante et complémentaire des études réalisées à large échelle sur la base de données spatiales imprécises (DUPOUEY *et al.*, 1999 ; METZGER *et al.*, 2006).

De plus, la méthodologie proposée a été mise en place à l'aide de données permettant une certaine reproductibilité de la démarche. En effet, plusieurs études récentes de l'estimation du

stock de carbone dans la végétation sont basées sur un jeu de données issue du LiDAR (ASNER, *et al.*, 2012). Ces données, bien que plus précises présentent l'inconvénient d'être très onéreuses, peu reproductibles et spatialement peu étendues. Cela en fait un outil peu adapté aux objectifs posés par le MEA ou les REDD. Au contraire, les données utilisées dans cet article, notamment en ce qui concerne les images LANDSAT TM ou du MNT ASTER, sont faciles à acquérir, en libre accès, couvrent de vastes espaces géographiques et présentent le dernier avantage de fournir des images à plusieurs dates, ce qui en fait un outil reproductible. La cartographie du stock de carbone telle que nous l'employons permet donc de retranscrire simplement des processus complexes par des proxys dérivés. Si l'on ne peut nier l'existence de marges d'erreur, ces proxys tiennent compte de la diversité paysagère. Cela en fait un outil essentiel pour les acteurs du territoire qui peuvent s'approprier le phénomène et aménager les territoires efficacement en s'adaptant aux spécificités locales, aux politiques publiques mises en place et aux réalités du terrain étudié. Cette complémentarité des échelles permettrait de répondre pleinement aux objectifs posés par le MEA ou les REDD en facilitant les transferts de connaissances et d'informations auprès des acteurs locaux et des gestionnaires internationaux en vue de la mise en œuvre d'une politique de gestion des ressources environnementales plus adaptée aux enjeux territoriaux concernés. En effet, la cartographie obtenue permet d'analyser les rapports Homme / Milieux puisqu'elle offre une meilleure compréhension de la répartition des stocks de carbone sur les territoires en cours de déforestation et peut être ainsi utilisée par les gestionnaires dans une optique d'aménagement plus durable de ces espaces. Cet aspect est important au Brésil, notamment dans les espaces situés dans l'arc de déforestation amazonien. Une meilleure compréhension de la dynamique des processus physiques influencés par l'anthropisation est essentielle afin d'assurer la réussite et la pérennisation d'actions internationales, notamment dans le cadre des programmes REDD et REDD+. Ainsi, le choix a été fait de cartographier des valeurs associées au SE « stockage de carbone », i.e. un état, (quantité stockée) et non pas le processus dynamique en lui-même (stockage de carbone), même si cette analyse des variations de ce stock dans l'espace permet de comprendre l'importance du phénomène de déstockage, depuis la forêt dense jusqu'au développement de pratiques anthropiques. Cet objectif répond à d'éventuelles applications dans les négociations internationales comme le REDD dont les processus de compensation proposés nécessitent une connaissance précise des stocks de carbone dans la végétation. Il est toutefois envisagé une cartographie du processus de stockage en étudiant le stock de carbone pour un même site à des dates différentes, ce qui permettrait de mettre en évidence la dynamique temporelle des stocks de carbone.

### **Références bibliographiques**

ACHARD F, EVA H D, MAYAUD P, STIBIG H-J and BELWARD A, 2004. Improved estimates of net carbon emissions from land cover change in the tropics for the 1990s Glob. Biogeochem. Cycles, 18 GB2008 doi:10.1029/2003GB002142

AKAIKE H., 1973. Information theory and an extension of the maximum likelihood principle. In : Petrov, B. N. and Csaki, B. F. (Eds), Second international symposium on information theory. Academia Kiado, Budapest, pp. 267-281

ARNAULD DE SARTRE X., BERDOULAY V., 2011. Des politiques territoriales durables ? Leçons d'Amazonie, Paris, Editions Quae, 2011, 158 p.

ASNER P.G., MASCARO J, MULLER-LANDAU C. H., VIEILLEDENT G. and VAUDRY R, 2012. A universal airborne LiDAR approach for tropical forest carbon mapping, *Oecologia*, Volume 168, Number 4, Pages 1147-1160

BOLFE et BATISTELLA, 2011. Carbon mapping of agroforestry systems in the Brazilian amazon through satellite images. Proceedings of the 25th International Cartographic Conference, Paris, 3-8 July 2011.

BREIMAN L., FRIEDMAN J., STONE C.J. and OLSEN R.A., 1984. Classification and regression Trees. Boca Raton ed., Monterey, USA, 159 p.

BROWN S., 1997. Estimating biomass and biomass change of tropical forests: A primer. FAO Forestry Paper: vii, 55 p.

BURKHARD B., KROLL F., MÜLLER F. and WINDHORST W., 2009. Landscapes' Capacities to Provide Ecosystem Services – a Concept for Land-Cover Based Assessments. *Landscape online*, 15 : 1-22.

CORNILLON P.A., GUYADER A., HUSSON F., JEGOU N., JOSSE J., KLOAREG M., MATZNER-LOBER E. et ROUVIERE L., 2010. Statistiques avec R, 2<sup>ème</sup> édition augmentée, Presses Universitaires de Rennes, France, 210 p.

CORNILLON P.A., MATZNER-LOBER E, 2011. Régression avec R, Springer, France, 242p.

DUPOUEY, J.L., PIGNARD, G., BADEAU, V., THIMONIER, A., DHÔTE, J.F., NEPVEU, G., BERGÈS, L., AUGUSTO, L., BELKACEM, S., NYS, C., 1999. Carbon stores and fluxes in French forests. *Comptes Rendus de l'Académie d'Agriculture de France*, 85, p. 293-310

FEARNSIDE P.M., 2005. Deforestation in Brazilian Amazonia: History, Rates, and Consequences. *Conservation biology*, 19(3), pp. 680-688.

EVANS, T.P., MANIRE, A., DE CASTRO, F., BRONDIZIO, E., MCCRACKEN, S., 2001. A dynamic model of household decision making and parcel level land cover change in the eastern Amazon. *Ecological Modelling* 143 (1-2), 95-113.

FUKUDA M., IEHARA T., MATSUMOTO M., 2012. Carbon stock estimates for sugi and hinoki forests in Japan. *Forest Ecology and Management* Volume 184, Issues 1–3, 3 October 2003, Pages 1–16.

GERWING J.J., 2002. Degradation of forests through logging and fire in the eastern Brazilian Amazon. *Forest Ecology and Management*, 157 : 153-165.

HIGUCHI N., SANTOS J., RIBEIRO R.J., MINETTE L. et BIOT Y., 1998. Biomassa da parte aerea da vegetação de floresta tropical úmida de terra-firme da Amazônia Brasileira. *Acta Amazonica*, 28 (2) : 153-165.

LECOINTE S., NYS C., WALTER C., FORGEARD F., HUET S., RECENA P., FOLLAIN S., 2006 - Estimation of carbon stocks in a beech forest (Fougères Forest – W. France): extrapolation from the plot to the whole forest. *Annals of Forest Science*, 63,1-11.

LIU W. et Yang Y., 2011. Parametric or nonparametric? A parametricness index for model selection, *The Annals of Statistics*, Volume 39, Number 4, 2074-2102

MALLOWS C.L., 1973. Some comments on Cp. *Technometrics* 15, 661-675

MARKEWITZ D., DAVIDSON E., MOUTINHO P. et NEPSTAD D., 2004. Nutrient loss and redistribution after forest clearing on a highly weathered soil in Amazonia. *Ecological Applications*, 14 : 177-199.

METZGER M.J., ROUNSEVELL M.D.A., ACOSTA-MICHLIK L., LEEMANS R. et SCHROTER D., 2006. The vulnerability of ecosystem services to land use change. *Agriculture, Ecosystems and Environment*, 114 : 69-85.

MONTGOMERY D.C., PECK E.A. and VINNING G.G., 2001. *Introduction to linear regression analysis*. John Wiley, 3 ed.

NAIDOO R., BALMFORD A., COSTANZA R., FISHER B., GREEN R.E., LEHNER B., MALCOLM T.R. and RICKETTS T.H., 2008. Global mapping of ecosystem services and conservation priorities. *PNAS*, 105 : 9495-9500.

NELSON W. B., MESQUITA R., PEREIRA L.G., DE SOUZA S.G.A., BATISTA G.T., COUTO L.B., 1998. Allometric regressions for improved estimate of secondary forest biomass in the central Amazon, *Forest Ecology and Management* 117, 149-167



OSZWALD J., ARNAULD DE SARTRE X., DECAENS T., GOND V., GRIMALDI M., LEFEBVRE A., DE ARAUJO FRETAS R.L., LINDOSO DE SOUZA S., MARICHAL R., VEIGA I., VELASQUEZ E. and LAVELLE P., 2012, Utilisation de la télédétection, de données socio-économiques et écologiques pour comprendre l'impact des dynamiques de l'occupation des sols à Pacajá (Brésil). *Revue Française de Photogrammétrie et de Télédétection*. Sous Presse.

OSZWALD J., GOND V., DOLEDEC S. et LAVELLE P., 2011. Identification d'indicateurs de changement d'occupation du sol pour le suivi des mosaïques paysagères. *Bois et Forêts des Tropiques*, 307 : 7-21.

PAVE A, FORNET G., 2010. Amazonie: Une aventure scientifique et humaine du CNRS, Galaade Editions, 156p

PONCE-HERNANDEZ R., 2007. A Modelling Framework for Addressing the Synergies between Global Conventions through Land Use Changes: Carbon Sequestration, Biodiversity Conservation, Prevention of Land Degradation and Food Security in Agricultural and Forested Lands in Developing Countries. FEEM Working Paper No. 30.2007. Available at SSRN: <http://ssrn.com/abstract=976600> or <http://dx.doi.org/10.2139/ssrn.976600>

RUTLEDGE D.T, DYMOND J., GREENHALH S., AUSSEUIL A.G., SINCLAIR R., HERZIG A., MORGAN F., ANDREW R. COLLIN, 2010. Mapping, Modelling and Managing Ecosystems Services in New Zealand, 8p

SALIMON I.C., PUTZ E.F., MENEZES-FILHO L., ANDERSON A, SILVEIRA M., BROWN F, OLIVEIRA L.C., 2011. Estimating state-wide biomass carbon stocks for a REDD plan in Acre, Brazil. *Forest Ecology and Management* 262, 555–560

SCHWARZ G., 1978. Estimating the dimension of a model. *The annals of statistics*, Vol 6, n°2, 461-464

SILVER W.L., OSTERTAG R. and LUGO A.E., 2000. The potential for Carbon Sequestration Through Reforestation of Abandoned Tropical Agricultural and Pasture Lands. *Restoration Ecology*, 8 : 394-407.

WAUTERS J.B., COUDERT S., GRALLIEN E., JONARD M., PONETTE M., 2007. Carbon stock in rubber tree plantations in Western Ghana and Mato Grosso (Brazil), *Forest Ecology and Management* 255, 2347–2361

## Annexe 1 : Estimation des paramètres du modèle complet en régression linéaire

	Estimate	Pr(> t )	
(Intercept)	226.0950	0.038629	*
Pente	2.5607	0.241878	
Altitude	-0.1508	0.795615	
Trajectoire2	20.2653	0.644370	
Trajectoire4	61.4265	0.289195	
Trajectoire5	67.6601	0.384596	
OS2	-178.3472	0.000933	***
OS3	-311.8502	0.000459	***
OS4	-320.4483	0.000389	***
OS5	-298.71	0.000377	***
OS6	-225.76	0.000341	***
Topographie2	47.8119	0.448717	
Topographie3	-13.017	0.787703	
Topographie4	2.9101	0.959920	
Deau200	-13.9826	0.831082	
Deau300	66.9702	0.389912	
Deau500	16.4048	0.810746	
Deau90000	35.0445	0.596721	

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Multiple R-squared: 0.6622  
p-value: 0.008035

Plusieurs paramètres sont non significativement différents de 0. Certaines variables n'apportent aucune information. Cela justifie la procédure de choix de variables.

## Annexe 2 : Estimation des paramètres du modèle choisi après sélection de variables en régression linéaire

	Estimate	Pr(> t )	
(Intercept)	281.15	6.43e-10	***
OS2	-186.54	3.35e-05	***
OS3	-278.36	5.75e-05	***
OS4	-281.02	5.05e-05	***
OS5	-313.72	1.04e-05	***
OS6	-232.87	7.22e-06	***
Deau300	99.95	0.0115	*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Multiple R-squared: 0.6037  
p-value: 4.313e-06