

# Déconstruire la spatialisation de services écosystémiques par la modélisation critique.

Par Solen Le Clec'h, Johan Oszwald, Simon Dufour, Michel Grimaldi, Nicolas Jégou et Matthieu Noucher

---

## Approche critique de la cartographie de services écosystémiques.

### Quantifier et spatialiser les objets de nature : un enjeu toujours actuel.

La carte, représentation partielle ou complète de la surface terrestre, sur un support plan, peut constituer un moyen de communication efficace. Elle peut, par exemple, faciliter le dialogue entre les chargés d'études scientifiques ou techniques et les responsables de l'aménagement du territoire (Long 1973) (Scarwell, Kergomard et Laganier 2008). De fait, les enjeux liés à la représentation cartographique sont depuis longtemps reconnus, notamment dans l'analyse de données relatives à l'état du milieu, à la biodiversité et aux processus environnementaux (Elhai 1968) (Allen et Starr 1982) (Turner et al. 2003). En France, par exemple, depuis la moitié du 20<sup>e</sup> siècle, les cartographies du milieu naturel établissent un lien entre la connaissance fondamentale des milieux et la gestion environnementale. Les différentes pratiques de cartographie de la végétation, quelles que soient leur école ou leur échelle, s'inscrivent en effet bien souvent dans des actions liées à des objectifs d'aménagement ou d'évaluation environnementale (Rey 2009). D'autres formes d'expression cartographique co-existent avec des cartographies de végétation, sous la forme de cartographies thématiques telles que les cartes phyto-écologiques, qui permettent de représenter spatialement les unités de végétation, c'est-à-dire la végétation en relation avec d'autres caractéristiques du milieu, telles que le climat, l'hydrologie ou la morphologie (Long 1973) (Amoros et al. 1982).

Depuis quelques années émergent également de nouvelles formes de fabriques cartographiques qui complètent, contournent ou concurrencent les référentiels géographiques conventionnels (Noucher 2017). Sous l'effet à la fois de la convergence des technologies de l'information et de la communication et des techniques de géolocalisation, on assiste, en effet, à une forte diversification des usages de l'information géographique. Cette expansion, notamment sur Internet, est largement perceptible dans les domaines de la gestion environnementale. Elle

accompagne bien souvent la montée en puissance de dynamiques collaboratives, en développant des cartographies participatives des milieux (Palsky 2013), à l'image de la démultiplication des projets de sciences citoyennes qui agrègent, sous la forme de cartes, des observations individuelles de l'environnement des « citoyens-capteurs » (Goodchild 2007). Les caractéristiques de ces données sont très diverses. Cela va de la plainte ponctuelle formulée par des citoyens inquiets de la qualité de leur environnement en mesurant des paramètres physico-chimiques à de larges bases de données alimentées par des réseaux d'informateurs distribués, relatives par exemple aux occurrences d'espèces, à leur distribution, à leurs habitats... (Georis-Creuseveau, Noucher et Gourmelon 2018). Les données recueillies sont à la fois qualitatives et quantitatives. Elles peuvent inclure (1) des éléments factuels, comme l'observation d'espèces spécifiques au sein d'une aire géographique, (2) des opinions ou des plaintes et (3) des mesures mobilisant des capteurs, tels que des GPS ou des pH mètre, par exemple (Gouveia et Fonseca 2008). Ces données sont généralement liées à un lieu et se caractérisent par une composante spatiale implicite (Newman et al. 2017).

Ainsi, les politiques de conservation de la biodiversité, grandes consommatrices d'informations géographiques, sont aujourd'hui confrontées à une montée en puissance des initiatives contributives (Amelot, Couderchet et Noucher 2017). La connaissance officielle qui peine à progresser est ainsi débordée par les démultiplications d'informations réunies et diffusées en dehors des circuits officiels. Ces différentes cartographies, aux méthodes souvent inédites, peuvent être centrées sur le milieu (par exemple sur la végétation) ou sur des composantes sociales (par exemple via des approches participatives), en écologie ou en géographie. Qu'elles soient à l'initiative de la sphère institutionnelle ou associative, elles possèdent une dimension opérationnelle : elles peuvent ainsi alimenter des cartes servant à définir des zonages réglementaires ou des contre-cartes visant à alimenter le débat public. Elles sont ainsi présentées comme un outil de la gestion environnementale et deviennent une ressource clé pour alimenter les politiques publiques (Gautreau et Noucher 2013), à l'image de toute forme de production de données (Fortier et Alphandery 2017). Ce faisant, elles participent à une forme de gouvernance informationnelle de l'environnement (Mol 2009) (Arpin, Charvolin et Fortier 2015) (Le Bourhis 2016) qui mérite d'être interrogée pour en comprendre les différents ressorts (politiques, techniques, éthiques). La cartographie des services écosystémiques s'inscrit clairement dans la lignée de ces approches, c'est-à-dire comme outil de la gestion environnementale approprié par des acteurs aux profils (et objectifs) de plus en plus variés.

## Mises en chiffre et mises en carte des services écosystémiques : des pratiques en fort développement.

Les politiques de l'environnement sont en grande partie fondées sur la définition de stratégies et la mise en œuvre d'outils qui proposent des indicateurs pour mettre en chiffres ou en cartes les objets de nature « à gérer » (Rumpala 1999) (Daniel 2010) (Guimont et Petitimbert 2017). Ces indicateurs s'appuient sur des concepts, tels que celui de services écosystémiques. Ces concepts permettent un aller-retour entre la production du savoir (la science) et l'application de ce savoir (politique). Ils tentent d'englober les systèmes écologiques et sociaux et leurs interrelations. Le concept de services écosystémiques, comme celui de biodiversité, fait partie des outils mixtes politico-scientifiques qui se situent à l'interface science-politique et qui visent / permettent une coproduction de ces deux domaines (Arnauld de Sartre et al. 2014). De fait, en l'espace de vingt ans, le concept de services écosystémiques s'est rapidement retrouvé au cœur des enjeux de la conservation et du développement.

Les services écosystémiques sont couramment définis comme « les bénéfices fournis par les écosystèmes, dont les services de prélèvement comme la nourriture, l'eau, le bois, les ressources génétiques ; les services de régulation tels que la régulation du climat, des inondations, des maladies, de la qualité de l'eau ou le traitement des déchets ; les services culturels comme les activités récréatives, les bénéfices liés à l'esthétique, au spirituel, et les services de soutien tels que la formation des sols, la pollinisation ou le cycle des nutriments » (Millenium Ecosystem Assessment 2005).

Spatialiser les services écosystémiques est possible car leur fourniture est spatialement explicite (Martinez-Harms et Balvanera 2012) et hétérogène (Fischer et al. 2008). En effet, leur agencement spatial est important pour leur productivité et leur qualité (Boyd et Banzhaf 2007). Les cartes de services écosystémiques, comme les autres formes d'expression cartographique, connaissent un réel engouement de la part des acteurs politiques. Cet engouement est notamment dû à l'apparition de moyens nouveaux (développement des outils de la géomatique, plateformes web) qui permettent d'accroître leur production et leur diffusion (Joliveau, Noucher et Roche 2013). Il peut s'expliquer par les intérêts essentiellement scientifiques et politiques que cette pratique soulève. La notion de services écosystémiques ayant d'ailleurs été globalement acceptée comme un moyen d'évaluation de l'état des écosystèmes à plusieurs échelles spatiales (Seppelt et al. 2011) (Crossman et al. 2013), sa cartographie devient légitimement un outil pertinent.

À l'instar de ceux de la notion elle-même, les enjeux liés à la spatialisation des services écosystémiques relèvent de trois ordres distincts : pédagogique, heuristique et politique – au sens de politique publique et de stratégie politique (Arnauld de Sartre et al. 2014). Historiquement, les cartographies de services écosystémiques répondaient davantage à un usage pédagogique (Costanza et al. 1997). En d'autres termes, la spatialisation des services écosystémiques est avant tout considérée comme une manière utile de représenter l'information (UNEP-WCMC 2011). La cartographie, en tant qu'outil visuel de communication, permet d'expliquer l'importance des services écosystémiques rendus sur le territoire (Maes et al. 2013). Les cartes de services écosystémiques ont ensuite, à l'image de la notion de services écosystémiques, glissé vers la dimension opérationnelle (Pesche 2011) (Hrabanski 2013), comme le programme européen MAES et sa déclinaison française EFESE, qui visent à cartographier les services écosystémiques pour l'ensemble des territoires concernés par le programme. À présent s'affirme l'aspect heuristique, avec le développement de réflexions scientifiques sur les pratiques cartographiques (Schulp et al. 2014).

## **Les difficultés et les enjeux liés à la cartographie d'objets de nature : le cas des services écosystémiques.**

La représentation cartographique présente une dimension technique forte. En effet, elle mobilise des savoirs, des outils et des méthodes spécifiques (Poidevin 1999). Elle relève également d'une dimension symbolique, puisqu'elle revêt une forme de pouvoir (Harley 1990) (Wood, Fels et Krygier 2010). La réalisation d'une carte passe, en effet, par une série de choix (échelle spatiale, données utilisées...) qui vont être influencés, voire imposés, par le créateur de la carte, ou son commanditaire, et vont dépendre de l'information qui doit être mise en évidence. Or, ces choix techniques impactent directement le résultat cartographique final et, de fait, son utilisation potentielle.

Les approches actuelles de cartographies de services écosystémiques s'appuient sur un grand nombre de méthodologies et de données extrêmement variées (Martinez-Harms et al. 2015). Les articles scientifiques qui présentent ces cartographies n'explicitent que rarement la motivation du choix de la méthodologie et des données, en relation avec l'objectif que remplit la carte (Nahuelhual et al. 2015). Or, spatialiser des services écosystémiques ou leurs indicateurs implique que le cartographe réalise de nombreux choix : sur les échelles spatiales d'analyse, sur les données d'entrée, sur les méthodes de traitements, etc. Ces choix ne sont pas sans conséquence (Grêt-Regamey et al. 2014) (Vorstius et Spray 2015). Pourtant, l'influence des méthodes de cartographie et des données utilisées sur la représentation spatiale de la fourniture

en services écosystémiques est encore largement méconnue (Kandziora, Burkhard et Müller 2013) (Lavorel et al. 2014). L'examen de la littérature montre qu'il existe encore des aspects à étudier, comme l'influence du choix de la méthode ou celui des données sur la carte résultant des analyses.

Il n'existe pas de consensus sur l'approche méthodologique à favoriser, ce qui se traduit par un nombre important d'approches différentes. La quasi-totalité des études s'appuie sur une des cinq approches suivantes : l'attribution d'une valeur de services écosystémiques par type d'occupation du sol (Eigenbrod et al. 2010) (Fisher et al. 2011), l'application de modèles écologiques communément acceptés fondés sur des équations ou des modèles mathématiques qui décrivent des processus biophysiques – eg. USLE (Universal Soil Loss Erosion) (Nelson et al. 2009) (Bai et al. 2012) –, l'application de modèles économiques qui reposent, par exemple, sur les coûts associés à la perte des services ou qui analysent les préférences et comportements des individus (Naidoo et Ricketts 2006) (Busch et al. 2012), l'application de méthodes statistiques qui estiment un indicateur de services à partir d'un jeu de données explicatif (Van Wijnen et al. 2012) (Czerepowicz, Case, et Doscher 2012), la mise en œuvre de méthodes participatives (Sherrouse, Clement, et Semmens 2011) (Darvill et Lindo 2015). Par ailleurs, un grand nombre de cartographies de services écosystémiques sont réalisées à partir d'une seule donnée, souvent l'occupation du sol (Costanza et al. 1997) (Kienast et al. 2009). Or, les services écosystémiques sont fournis par des processus biophysiques ou des fonctions écologiques dont le fonctionnement est très diversifié, voire complexe. En effet, les services dépendent de multiples interactions. De fait, il paraît essentiel de prendre en compte une grande variété d'informations lors de l'application d'une méthode statistique, afin de réduire le phénomène étudié à un objet quantifiable, c'est-à-dire un indicateur de services écosystémiques. Les spatialisations fondées sur la seule donnée d'occupation du sol ont par conséquent été dénoncées comme simplificatrices (Eigenbrod et al. 2010). Actuellement, les données utilisées se diversifient grâce à l'utilisation, par exemple, des données de topographie, de climat ou encore de répartition des types de sol (Egoh et al. 2009) (Nedkov et Burkhard 2012). La nécessité d'ajouter de nouvelles variables, et notamment des données relatives à l'évolution temporelle du paysage, est d'ailleurs mise en avant (Ministère de l'Écologie, du Développement Durable et de l'Énergie 2009).

## **Les différents registres de la critique des métrologies de la Nature.**

Fort de ce triple constat (variété des approches modélisatrices, démultiplication des sources de données, absence de consensus sur la méthode) de nombreuses voix se sont élevées contre les approximations des mises en cartes de la nature. Nous relevons ici trois registres de critique et en

proposons une quatrième.

Un premier registre renvoie à la cartographie critique, déjà évoquée précédemment autour du travail fondateur de Brian J. Harley sur le « pouvoir des cartes » (Harley 1990). Prônant des analyses qui passent par la « déconstruction des cartes », il a ouvert la voie à de nombreux travaux en géographie et plus généralement en sciences sociales, démontrant à la fois le « pouvoir externe » et le « pouvoir interne » de la carte, qui font référence, pour le premier, à l'intentionnalité sociale et, pour le second, aux effets d'iconisation du média cartographique. Le « pouvoir externe » est ainsi le pouvoir du cartographe et de son commanditaire, qui sont les seuls à faire des choix sur ce qui va être montré ou caché, sur ce qui va être nommé ou non, sur la sémiologie graphique, etc. Les analyses multiples de l'intentionnalité sociale de la carte ont permis de mettre en exergue son caractère fondamentalement rhétorique. De nombreuses études historiques (Casti 2001) (Blais 2014) démontrent ainsi que les cartes ne constituent pas des relevés passifs d'objets géographiques, mais sont au contraire chargées de valeurs et influencées par de multiples facteurs (classe sociale, genre, religion, ethnicité, etc.). Le « pouvoir interne » renvoie, quant à lui, aux effets d'iconisation de la carte. Il relève l'impact des représentations cartographiques sur la pensée des acteurs qui les utilisent, du fait de la performativité de ce média (Craib 2000). Ainsi, non seulement les cartes ne sont pas neutres, mais elles agissent sur le changement social en véhiculant des catégories particulières d'analyse du monde. La cartographie critique et, en suivant, les SIG critiques (Pickles 2004) peuvent alors être mobilisées pour « donner à voir » des formes différenciées d'appropriation de l'espace et des jeux de pouvoir, à partir de cas d'étude. Cependant, le caractère récent de la spatialisation des services écosystémiques rend l'observation de situations de terrain encore limitée. Par ailleurs, face au développement de la géographie numérique qu'il a peu étudiée, Harley était méfiant et considérait que les méthodes assistées par ordinateur de la fin des années 1980 « rendent la rhétorique scientiste des cartographes encore plus marquée » (Harley 1990, p. 1). L'opacité des méthodes développées par les nouvelles fabriques cartographiques – et la spatialisation des services écosystémiques en est un bon exemple – complexifie aujourd'hui l'analyse des intentionnalités associées aux nouvelles métrologies de la nature.

Vers la fin des années 1990, les SIG critiques deviennent une sous-discipline de la géographie numérique (Desbois 2015), qui voit se développer différentes branches autour, par exemple, des SIG participatifs destinés à introduire les techniques géomatiques dans des actions de démocratie participative (Rambaldi 2005). Cette deuxième approche critique vise alors à « inverser » le « pouvoir des cartes », en donnant potentiellement la voix à diverses minorités pour les accompagner dans la production de contre-cartes. Dans le domaine des services

écosystémiques, de multiples initiatives autour de la spatialisation d'indicateurs des services culturels ont été expérimentées (Ribeiro et Ribeiro 2016). Cependant, de nombreux exemples illustrent les difficultés soulevées par les approches participatives, notamment la timidité, voire la paralysie des participants face à la complexité de la notion de services écosystémiques (Olszaska et al. 2015). L'effet dialogique de ces dispositifs participatifs a déjà été souligné dans d'autres contextes (Chambers 2006) (Hirt 2009). D'une part, l'expansion des usages de la carte, au-delà des cartographes patentés, contribue à renforcer les capacités des non-experts de la cartographie de se réapproprier le pouvoir de se définir et de se représenter par eux-mêmes. Elle renvoie ainsi à un processus d'émancipation et d'*empowerment*. Face à la cartographie institutionnelle, les contre-cartographies proposent des contre-expertises en s'appuyant sur des « savoirs locaux ». Les cartes produites cherchent alors à faire bouger les modèles dominants en remobilisant leur formalisme. Mais pour être visible dans le débat public et exister parmi la myriade d'indicateurs produits, il est nécessaire d'utiliser les « outils » reconnus par les autorités en place. Le deuxième effet de cette dialogique est donc de participer à un processus d'assimilation culturelle par la technique, qui peut alors renforcer les jeux de domination. La dialogique *empowerment* / assimilation de ces métrologies alternatives suscite donc de réelles interrogations quant à leurs impacts, car elles sont porteuses d'une promesse paradoxale : elles offrent une diversité de représentations de l'espace, tout en mettant en visibilité les contenus les plus formellement uniformes.

Un troisième registre de critique rejette toute mise en chiffre de la nature. Cette critique se fonde sur la complexité des systèmes socio-environnementaux étudiés. Ces approches critiques ne cherchent pas à pénétrer la complexité des protocoles de fabrique des indicateurs de services écosystémiques. En se positionnant sur un registre théorique, sans analyse des processus socio-techniques qui sont à l'œuvre, ces travaux ont alors pu être tentés par des postures spéculatives quelque peu déconnectées du terrain (Costanza et al. 1997) (Troy et Wilson 2006).

Aussi, un quatrième registre voit progressivement le jour. Il s'appuie sur des méthodes exploratoires qui favorisent l'émergence des analyses par une immersion dans les données et les modèles, pour décrypter la fabrique de l'information et en révéler toutes les marges d'incertitude qui y sont associées. Cette approche s'inscrit dans une logique de « modélisation critique » qui, telle que définie par Nicolas Bouleau (2014), vise à favoriser la démultiplication des « contre-modélisations », pour non seulement critiquer les modèles mais également contribuer à amender et enrichir l'appréhension des problèmes sous un jour pluraliste. Partant du constat qu'un modèle élaboré avec soin est très difficile à « défaire » car ses hypothèses implicites et ses non-dits sont souvent complexes à déconstruire, voire tout simplement à déceler, la modélisation critique

propose de concevoir plusieurs modèles pour faire émerger, autour de chaque phénomène étudié, des « co-vérités ». Celles-ci ne révéleraient pas seulement des parties cachées du réel, elles modifieraient les enjeux, les validations à entreprendre et, *in fine*, l'établissement des légitimités. Ce faisant, les co-vérités territoriales qui émergent d'une telle approche peuvent agir comme un puissant ressort de participation : en montrant que, malgré la complexité technique, il peut y avoir débat, on favorise le débat (Bobbio et Melé 2015).

C'est en s'inspirant de ce quatrième registre de critique que l'approche employée dans le cadre de cet article [\[1\]](#) cherche à approcher la spatialisation des indicateurs de services écosystémiques sous un jour pluraliste. La démarche repose alors sur l'application de méthodes statistiques qui permettent de prédire des valeurs de services écosystémiques sur un territoire, à partir de différentes données de terrain. Bien qu'encore peu développée, cette démarche présente deux avantages majeurs, particulièrement intéressants pour, d'un point de vue tant opérationnel que pédagogique, démontrer la diversité du champ des possibles. Par une procédure de choix de variables, les méthodes statistiques peuvent permettre de comprendre quels sont les facteurs qui influencent la fourniture de services écosystémiques à partir de la compréhension des processus biophysiques et des conditions environnementales. De plus, elles permettent d'estimer le degré de fiabilité des estimations, via l'utilisation d'indices tels que le coefficient de détermination ( $R^2$ ). En d'autres termes, elles offrent la possibilité d'estimer la pertinence de chaque variable (occupation du sol, topographie...) à répondre à la spatialisation d'un indicateur de services écosystémiques. D'un point de vue opérationnel, cette approche permet de fournir au gestionnaire un niveau de confiance en la carte produite, évitant le leurre de la solution « universelle ».

En s'inspirant de la modélisation critique, cet article vise donc à illustrer 1) une démarche alternative aux spatialisations actuelles de services écosystémiques et 2) l'impact de certains choix méthodologiques induits par la pratique cartographique sur la métrologie des objets de la Nature, à partir de l'exemple de la spatialisation des indicateurs de services écosystémiques. Pour ce faire, nous nous intéressons, de façon privilégiée, à l'impact du choix et de la précision des données d'entrée, ce qui constitue une première approche pour tendre à une appréhension pluraliste des processus socio-écologiques étudiés.

## Matériel et méthodes.

### Approche générale.

Afin d'analyser l'impact du choix et de la précision des données d'entrée sur notre capacité à fournir une cartographie d'indicateur de services écosystémiques fiable, une méthodologie en deux étapes a été mise en place :

1. La première vise à appliquer une méthode de régression pour modéliser trois indicateurs de services écosystémiques à partir d'un jeu de données variées obtenues par traitement d'images satellite. En d'autres termes, nous cherchons à estimer pour l'ensemble des pixels de chaque zone d'étude des valeurs pour ces trois indicateurs de services écosystémiques. Or, ces informations ne sont connues qu'au niveau des surfaces échantillonnées où elles ont été mesurées *in situ*. Par conséquent, pour chaque indicateur de services écosystémiques, nous appliquons deux méthodes de régression : l'arbre de régression et le modèle linéaire. La méthode de régression qui présente la meilleure qualité prédictive est sélectionnée de manière objective, par procédure de validation croisée, pour chaque indicateur. Cette méthode est ensuite utilisée pour prédire des valeurs d'indicateurs de services écosystémiques sur la surface des territoires étudiés, permettant ainsi de réaliser une cartographie.
2. La seconde étape constitue un retour sur l'expérience de l'étape précédente. Alors que de nombreuses cartographies d'indicateurs de services écosystémiques se fondent sur une donnée simple, l'objectif de cette seconde étape est d'évaluer l'apport d'un jeu de données variées (en nombre et en thématiques traitées) pour modéliser des indicateurs de services écosystémiques. En d'autres termes, les analyses menées visent à évaluer l'impact du choix et de la précision des données d'entrée sur notre capacité à fournir une cartographie d'indicateur de services écosystémiques fiable. Pour ce faire, deux analyses distinctes ont été mises en œuvre :
3. Tout d'abord, nous cherchons à évaluer l'impact du choix et de la précision des données d'entrée. En d'autres termes, cela revient à déterminer l'apport d'un jeu de données variées pour modéliser des indicateurs de services écosystémiques. En effet, beaucoup d'approches cartographiques des services écosystémiques se basent sur la seule donnée d'occupation du sol (Troy et Wilson 2006) (Swetnam et al. 2011) (Scolozzi, Morri et Santolini 2012). Il semble essentiel d'évaluer la plus-value d'approches basées sur des données plus variées et de déterminer un seuil de données nécessaires, tout en garantissant une facilité d'acquisition. Pour répondre à cet objectif, nous nous basons sur des modèles de régression linéaire qui mettent en relation chacun des indicateurs des services écosystémiques étudiés et des données de télédétection. Nous cherchons ainsi à expliquer successivement chaque indicateur par des données accessibles par télédétection, en

faisant varier ces dernières. La comparaison des différentes régressions par validation croisée permet d'estimer de manière objective le gain d'information que provoque l'ajout de données et qui conditionne la fiabilité de la cartographie d'indicateurs de services écosystémiques résultant de l'application de ces régressions.

4. Dans un second temps, en nous focalisant sur le cas des stocks de carbone du sol, une analyse plus fine a été mise en œuvre à partir de données plus diversifiées, en partie non accessibles par télédétection. Des régressions ont ainsi été implémentées en considérant un grand nombre de variables explicatives, à la fois issues de la télédétection et d'enquêtes de terrain.

## Sites d'étude.

Pour répondre à cet objectif, des cartographies individuelles d'indicateurs de services écosystémiques ont été réalisées à l'échelle locale pour trois sites, Maçaranduba, Pacajá et Palmares II, appartenant tous trois à l'État du Pará (Figure 1).

L'État du Pará est un État brésilien de 1,3 millions de km<sup>2</sup>, appartenant à la région d'Amazonie. Il présente des taux de déforestation parmi les plus élevés au Brésil<sup>[2]</sup>. La déforestation se fait essentiellement au profit de terres agricoles pour l'élevage bovin. Or, les dynamiques de déforestation affectent de manière négative les services écosystémiques, jusqu'à générer des changements qui altèrent le climat, la biodiversité, les sources d'eau potable et favorisent l'érosion du sol (Fearnside 2005) (Spracklen, Arnold et Taylor 2012).

Les trois sites locaux connaissent des dynamiques de déforestation, mais présentent des situations contrastées (Albaladejo et al. 1996). Le site de Maçaranduba couvre environ 220 km<sup>2</sup>. Il est localisé sur le municipe (commune) de Nova Ipixuna, en bordure de fleuve. Le déboisement du site a commencé dans les années 1970. Au début des années 2000, un type d'organisation sociale comparable aux réserves agro-extractivistes a été mis en place, avec, notamment, propriété collective de la terre, attribution de droits d'usage sur des parcelles et limitation des autorisations de déboisements pour l'implantation de pâturages. Les exploitants mènent à la fois des activités agricoles et des activités qui valorisent des ressources forestières, notamment en production secondaire. De fait, bien que situé dans une zone largement déboisée, ce site possède une couverture forestière relativement bien préservée, avec quelques patches de forêt conservée. Le site de Pacajá (175 km<sup>2</sup>), localisé sur la commune de Pacajá, à 60 km environ du centre urbain, se trouve sur une des pistes secondaires perpendiculaires à la Transamazonienne, appelées « arête de poisson », sur le Travessão sud du kilomètre 338. Ce site a connu une

colonisation spontanée intense, à partir de 1990. Enfin, Palmares II est un « assentamento » (localité de la Réforme Agraire) de 158 km<sup>2</sup>, de la commune de Parauapebas. Ce site est emblématique de la réforme agraire. Sa création a été décidée suite à des luttes entre les travailleurs du « Mouvement Sans Terre » (MST), le Gouvernement Fédéral et la Vale, propriétaire de l'exploitation minière de Carajas (Bringel 2006). L'expulsion du « fazendeiro », en 1996, a provoqué un émiettement de la zone en une multitude de parcelles.

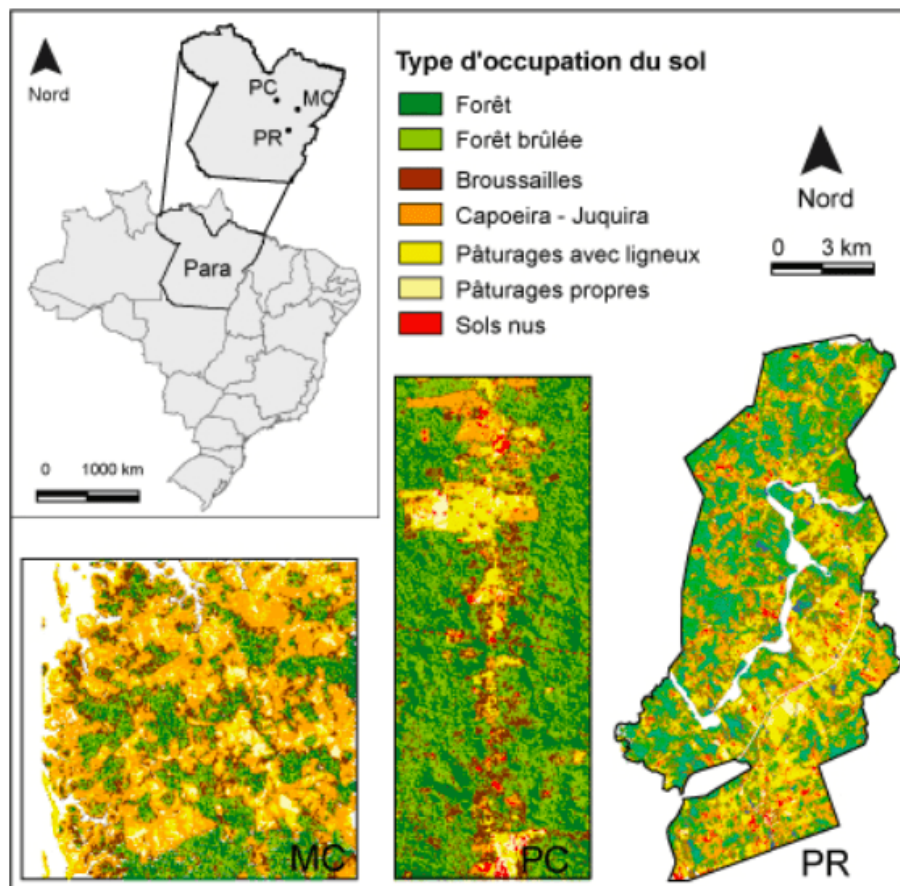


Figure 1 : Localisation de l'État du Pará et des trois sites d'études. Ces trois sites sont présentés sur fond de leur occupation du sol (2007). PC : Pacajá ; MC Maçaranduba et PR : Palmares II.

## Données.

### Variables à expliquer : indicateurs de services écosystémiques mesurés sur le terrain.

Les services écosystémiques sont évalués par l'intermédiaire d'indicateurs simples de processus biophysiques. La cartographie résultante repose donc sur des indicateurs quantifiables et

spatialisables, accessibles par télédétection.

Trois indicateurs de services écosystémiques ont été choisis, pour leur complémentarité et leur pertinence dans le contexte des fronts pionniers amazoniens. Ces indicateurs sont liés à plusieurs services écosystémiques : la régulation du climat, via les indicateurs de stocks de carbone de la biomasse végétale et le stock de carbone du sol, la régulation du cycle de l'eau et le contrôle de l'érosion du sol, via la capacité du sol à permettre l'infiltration de l'eau ; et le service de support à la production primaire, via les stocks de carbone du sol qui contribue à sa fertilité, particulièrement en milieu tropical (Grimaldi et al. 2014).

Le stockage du carbone dans la végétation est considéré classiquement comme un service de régulation (au sens du *Millenium Ecosystem Assessment* – MEA) du climat. Par son activité de photosynthèse, la végétation chlorophyllienne absorbe une partie du CO<sub>2</sub> de l'atmosphère et joue ainsi un rôle important dans la régulation du climat (Brown 1997), et ce particulièrement dans les forêts tropicales (Silver, Ostertag, et Lugo 2000).

L'infiltration de l'eau dans le sol est le processus par lequel le sol absorbe l'eau des précipitations. Elle est caractérisée par la hauteur d'eau qui pénètre dans un sol pendant un intervalle de temps déterminé. Elle peut être considérée comme un des indicateurs de régulation du cycle de l'eau.

Le service de production primaire est étudié à partir des stocks de carbone du sol, puisque ces stocks ont une influence sur la fertilité chimique du sol. L'importance de la fertilité des sols dans la fourniture de services écosystémiques est peu mise en avant dans les études scientifiques sur les services écosystémiques, mais a été soulignée par le MEA (2005), *The Economics of Ecosystems and Biodiversity* (Sukhdev et al. 2010) et la Convention on Biological Diversity (Walpole et al. 2011).

Les informations sur les services écosystémiques ont été recueillies pour 45 points situés dans neuf des exploitations (à raison de cinq par exploitation) présentes sur chaque localité, lors de campagnes terrain effectuées dans le cadre des programmes de recherche AMAZ-BD et AMAZ-ES (dont le coordinateur est Patrick Lavelle), entre 2007 et 2010, financés par l'ANR[3].

## **Variables explicatives.**

### **Données géographiques obtenues par télédétection.**

Les informations spatiales utilisées dans les modèles statistiques permettant de cartographier le stock de carbone dans la végétation et l'infiltration de l'eau dans le sol sont de deux types (table

1). Les premières sont des données relatives à l'occupation du sol (OS) : une classification de l'OS de 2007, une trajectoire historique issue d'une analyse diachronique (Oszwald et al. 2011) et deux indices de végétation[4]. Ces informations ont été obtenues par traitement d'images Landsat TM (30 x 30m). Les autres données sont relatives à la topographie (altitude en mètres, pente en pourcentage, synthèse de la topographie et distance au réseau hydrographique) et proviennent du traitement du MNT ASTER (Modèle Numérique de Terrain – 30 x 30m). En outre, une information sur le site d'appartenance des mesures a été ajoutée aux régressions, afin de prendre en compte les variations dues aux différences entre les sites. Cette donnée qualifie chaque surface de mesure des indicateurs de services écosystémiques et chaque pixel par son site d'appartenance : Maçaranduba, Pacajá et Palmares II.

| Données                                    | Source                            |
|--------------------------------------------|-----------------------------------|
| Occupation du sol                          | Landsat TM – Oszwald et al., 2011 |
| Trajectoire historique d'occupation du sol | Landsat TM – Oszwald et al., 2011 |
| NDVI                                       | Landsat TM                        |
| NDWI                                       | Landsat TM                        |
| Altitude (m)                               | MNT Aster                         |
| Pente (%)                                  | MNT Aster                         |
| Distance à l'eau                           | MNT Aster                         |
| Topographie                                | MNT Aster                         |
| Site                                       |                                   |

Table 1 : Données issues de la télédétection utilisées pour réaliser des régressions.

Les régressions linéaires mises en place permettent de tester l'influence de ces informations sur les variations des deux indicateurs de services écosystémiques et de sélectionner uniquement les variables influentes afin de réaliser la prédiction des valeurs de services écosystémiques.

#### Données obtenues lors des campagnes de terrain.

Lors de la troisième étape, les jeux de données présentés préalablement ont été complétés par des informations obtenues lors des campagnes terrain : le type de sol ainsi que plusieurs informations sur l'histoire du paysage (table 2). Plusieurs variables caractérisent chacune des 135 surfaces échantillonnées. Ces variables sont liées à l'historicité de l'occupation de son sol (par exemple, le nombre de brûlis qu'elle a subi ou le nombre de jachères qui y ont eu lieu). Ces informations n'ont été utilisées que lors de la construction de régressions qui ne seront pas utilisées par la suite pour réaliser des prévisions. En effet, ces données étant exhaustives

(connues seulement au niveau des zones de mesure et inconnues pour l'ensemble des sites), elles ne peuvent pas être utilisées dans un but prédictif. Par ailleurs, le type de sol a également été utilisé. Cette donnée a été échantillonnée sur le terrain en 2007 et est connue pour l'ensemble des 135 surfaces échantillonnées. Étant inconnue pour l'ensemble des pixels des sites, elle sert ici dans l'élaboration d'une régression qui ne sera pas utilisée pour réaliser une cartographie.

| Données                                                                                                                                                                                             | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Nombre de brûlis<br>Date du 1 <sup>er</sup> brûlis<br>Date du dernier brûlis<br>Date de la 1 <sup>re</sup> jachère<br>Date de la dernière jachère<br>Nombre de types d'occupation du sol différents | Nombre de brûlis effectués au niveau de l'espace de mesure des services écosystémiques<br>Nombre d'années écoulées depuis le premier brûlis<br>Nombre d'années écoulées depuis le dernier brûlis<br>Nombre d'années écoulées depuis la première mise en jachère<br>Nombre d'années écoulées depuis la dernière mise en jachère<br>Nombre d'occupation du sol depuis la déforestation (en considérant que toute la région est historiquement sous couvert forestier). Une valeur de 0 correspond à une surface forestière.                                                |
| « Type de sol »<br>4 modalités                                                                                                                                                                      | Ferralsols (L) : sols profonds bien drainés (altérite à plus de 1,2 m de profondeur), car très poreux et très perméables,<br>Acrisols (A) : sols assez profonds (altérite à moins de 1,2 m de profondeur), bien drainés sur les versants,<br>Cambisols (C) : peu profonds (altérite à moins de 0,6 m de profondeur), bien drainés sur les versants,<br>Sols hydromorphes (H) : peu profonds et gorgés d'eau, présents essentiellement dans des dépressions ou des terrains très marécageux et marqués morphologiquement par la stagnation de l'eau en saison des pluies. |

Table 2 : Données issues d'enquêtes ou de relevés de terrain utilisées pour réaliser des régressions qui n'ont aucune visée prédictive.

## Analyses statistiques.

### Sélection de la méthodologie statistique.

Dans un premier temps, nous avons sélectionné la méthode de régression la plus pertinente pour cartographier chaque indicateur de services écosystémiques. Pour ce faire, les données obtenues par télédétection (relatives à l'occupation du sol ou à la topographie) et les données *in situ* (indicateurs de services), toutes les deux connues au niveau des 135 surfaces d'échantillonnage, ont été mises en relation à l'aide d'une méthode de régression. Ces méthodes de régression permettent, par la suite, d'extrapoler les indicateurs de services écosystémiques à l'ensemble de la localité, à partir de quelques surfaces échantillonnées.

Dans la mesure où il s'agit de modéliser une variable quantitative (indicateur de services) à partir de variables qualitatives (ex. occupation du sol) et quantitatives (ex. altitude), deux méthodes ont été utilisées : l'arbre de régression (Algorithme Classification And Regression Trees – CART) et la régression linéaire multiple (Cornillon et al. 2010).

Pour chaque méthode, une procédure de sélection de variables a été implémentée. Elle permet d'identifier les variables explicatives influençant la variable à expliquer. Cela nous permet d'expliquer les variations de chacun des indicateurs de services écosystémiques à partir des données sélectionnées et ainsi de comprendre les facteurs qui influencent les processus écologiques mis en jeu. Les deux méthodes ont ensuite été comparées par validation croisée. Pour ce faire, l'ensemble des données a été divisé en un échantillon d'apprentissage (90 % des individus statistiques), sur lequel chaque régression est construite, et d'un échantillon test (10 % des individus statistiques) sur lequel la régression est appliquée pour faire de la prévision (Refaeilzadeh, Tang et Liu 2009). Une erreur de prévision (*mean square error*), mesurant l'écart entre les valeurs prédites sur l'échantillon test et les valeurs observées sur le terrain, est calculée pour chaque régression. Celle présentant le score de validation croisée le plus faible est ensuite utilisée pour le reste des analyses.

## **Évaluation de l'apport d'un jeu de données variées pour modéliser des indicateurs de services écosystémiques.**

### **L'apport de la prise en compte de données variées.**

Après avoir déterminé pour chaque indicateur la méthode de régression la plus pertinente, nous avons évalué l'apport d'un jeu de données variées (en nombre et en thématiques traitées) pour modéliser des indicateurs de services écosystémiques. Pour ce faire, nous avons tenté d'évaluer la plus-value potentielle d'approches fondées sur des données plus variées et nous avons aussi tenté de déterminer un seuil de données nécessaires pour obtenir une cartographie fiable en limitant les temps et coûts supplémentaires liés à la récolte et aux traitements des données. Pour comprendre en quelle mesure l'ajout d'informations spatialisées peut améliorer significativement la prévision statistique, deux analyses ont été menées.

Tout d'abord, nous avons fondé notre analyse sur l'étude de trois indicateurs de services écosystémiques présentés : les stocks de carbone de la végétation, les stocks de carbone du sol et l'infiltration de l'eau dans le sol. En employant la méthode de régression la plus performante, sélectionnée dans la première étape (arbre de régression pour les stocks de carbone de la

végétation et modèle linéaire pour les deux autres indicateurs), six régressions ont été construites pour chacun des indicateurs – chaque régression présentant un degré de complexité supérieur au précédent :

- La première régression met en relation chacun des deux indicateurs avec le NDVI.
- La seconde met en relation les indicateurs et une classification d'occupation du sol en deux modalités, présentant une opposition forêt / non-forêt.
- La troisième régression correspond à la mise en relation de chacun des indicateurs de services écosystémiques avec une classification d'occupation du sol en six modalités.
- La quatrième lie les indicateurs de services écosystémiques avec la classification d'occupation du sol (six modalités) et la trajectoire historique d'occupation du sol.
- À cette dernière régression, la donnée renseignant la topographie est ensuite ajoutée.
- Enfin, la dernière analyse met en relation les informations de deux indicateurs avec les données d'occupation du sol, des données dérivées du MNT ASTER et la variable liée au site d'étude.

Au cours des analyses multivariées, une procédure de choix de variables est systématiquement effectuée. Pour chaque indicateur, les six régressions sont comparées par validation croisée, afin de sélectionner celle possédant la capacité prédictive la plus élevée.

#### **Une régression à visée explicative.**

Dans un second temps, nous avons cherché à identifier des facteurs de contrôle de la fourniture en services écosystémiques. Une méthode de régression permettant d'ajuster au mieux la variable  $y$  à expliquer a été mise en œuvre. Notre analyse est fondée sur le cas d'étude des stocks de carbone du sol, car celui-ci est difficile à modéliser à partir des seules données issues de la télédétection. Des données télédétectées et des données issues de campagnes terrain (observations et échantillonnage pour des analyses au laboratoire) ont été utilisées. Ces dernières informations n'étant connues qu'au niveau des surfaces d'échantillonnage, cette régression a pour ambition d'expliquer au mieux le processus à partir de données relatives à la nature du sol et aux informations liées à l'emprise historique des hommes sur le paysage, sans qu'une prévision de valeurs ne puisse avoir lieu. Si cela conditionne notre capacité à réaliser une cartographie de l'indicateur de services écosystémiques, cela nous permet de mieux appréhender les limites de la cartographie construite à partir des seules données télédétectées. Cette analyse vise à comprendre dans quelle mesure l'ajout d'informations exhaustives aux régressions peut permettre d'améliorer la compréhension des facteurs influençant les indicateurs de services

écosystémiques, sans chercher à prédire de nouvelles valeurs.

## Résultats.

### Sélection de la méthodologie statistique.

De manière générale, les méthodes statistiques employées apparaissent comme pertinentes pour expliquer et prédire les indicateurs de services écosystémiques à partir des données télédéctées (coefficient de détermination :  $R^2 = 0,75$  pour les stocks de carbone de la végétation – table 3). Bien que les coefficients de détermination soient ici présentés, les méthodes ont été comparées et les meilleures ont été sélectionnées à partir des scores de validation croisée. En effet, ces derniers sont plus fiables que le coefficient de détermination, dont la valeur tend à augmenter mécaniquement avec le nombre de variables. La plupart des prévisions sont donc effectuées à partir du modèle linéaire : les stocks de carbone de la végétation et l'infiltration de l'eau dans le sol. Les stocks de carbone du sol ont été prédits à partir d'un arbre de régression.

| Indicateur modélisé               | Arbre de décision |      | Modèle linéaire |      | Méthode retenue     |
|-----------------------------------|-------------------|------|-----------------|------|---------------------|
|                                   | R2                | VC   | R2              | VC   |                     |
| Carbone de la végétation          | 0,75              | 1    | 0,75            | 0,92 | Régression linéaire |
| Infiltration de l'eau dans le sol | 0,66              | 1    | 0,57            | 0,86 | Régression linéaire |
| Stocks de carbone du sol          | 0,52              | 0,93 | 0,47            | 1    | Arbre de décision   |

Table 3 : Coefficient de détermination ( $R^2$ ) et score de validation croisée (VC – erreur de prévision cumulée) standardisé (pour chaque indicateur, chacun des scores a été divisé par la valeur maximale) de chaque régression pour les trois indicateurs de services écosystémiques.

D'un indicateur à l'autre, les coefficients de détermination associés à l'implémentation des deux méthodes statistiques peuvent être très différents. La régression implémentée pour les stocks de carbone du sol possède ainsi une capacité prédictive très faible. Cette faible capacité prédictive illustre les limites de l'utilisation de données télédéctées pour prédire des services écosystémiques. En effet, les variations de certains indicateurs, comme les stocks de carbone de la végétation, peuvent essentiellement, voire exclusivement, être expliquées par des variables liées aux propriétés modifiables des écosystèmes, qui peuvent être en partie dérivées d'images

satellite. Toutefois, les variations d'autres indicateurs, tels que ceux liés aux processus biophysiques du sol, relèvent des propriétés inhérentes aux milieux, qu'il est plus difficile d'obtenir par télédétection (Dominati, Patterson et Mackay 2010).

Outre la variabilité de leur capacité à prédire, les méthodes se différencient par les données explicatives retenues, en matière de nombre et d'identité (table 4). De manière générale, le nombre de variables retenues est très élevé, ce qui souligne la complexité des phénomènes étudiés. Ce résultat montre bien la nécessité de posséder un jeu de données riche et d'adapter celui-ci à l'indicateur de services écosystémiques étudié. Quatre des six régressions ayant été mises en œuvre retiennent la variable liée au site (appartenance à Maçaranduba, Pacajá ou Palmares II). Ce résultat indique ainsi qu'il existe un fort effet de contexte géographique et/ou historique. En effet, chaque site est caractérisé par des politiques publiques, une historicité de la colonisation et une mise en valeur du territoire spatialement différente. Ces dernières influencent donc grandement la fourniture de services écosystémiques. Pour un même indicateur, l'arbre de décision et la régression linéaire peuvent retenir des variables très différentes et présenter des capacités prédictives très contrastées. Pour certains indicateurs, comme les stocks de carbone de la végétation, les deux méthodes mettent en avant le rôle des mêmes facteurs géographiques et environnementaux. Néanmoins, dans d'autres cas, comme pour l'infiltration de l'eau dans le sol, les variables retenues sont relativement différentes d'un modèle à un autre.

| Indicateur modélisé                | Arbre de décision                                                                          | Modèle linéaire                                         | Méthode retenue     |
|------------------------------------|--------------------------------------------------------------------------------------------|---------------------------------------------------------|---------------------|
| Stocks de carbone de la végétation | OS et site                                                                                 | OS et site                                              | Régression linéaire |
| Infiltration de l'eau dans le sol  | Trajectoire d'occupation du sol, NDWI, OS, altitude, distance à l'eau, topographie et NDVI | Trajectoire d'occupation du sol et OS                   | Régression linéaire |
| Stocks de carbone du sol           | Site, distance à l'eau, trajectoire d'OS, topographie, NDWI et OS                          | Site, topographie, trajectoire d'OS et distance à l'eau | Arbre de décision   |

Indicateur modélisé    Arbre de décision    Modèle linéaire    Méthode retenue  
Stocks de carbone de la végétation    OS et site    OS et site    Régression linéaire  
Infiltration de l'eau dans le sol    Trajectoire d'occupation du sol, NDWI, OS, altitude, distance à l'eau, topographie et NDVI    Trajectoire d'occupation du sol et OS    Régression linéaire  
Stocks de carbone du sol    Site, distance à l'eau, trajectoire d'OS, topographie, NDWI et OS    Site, topographie, trajectoire d'OS et distance à l'eau    Arbre de décision

topographie, trajectoire d'OS et distance à l'eau Arbre de décision

Table 4 : Variables retenues pour chaque méthode statistique mise en œuvre pour les trois indicateurs de services écosystémiques. La variable en gras correspond à la variable la plus influente. OS = occupation du sol.

Évaluation de l'apport d'un jeu de données variées.

L'apport de la prise en compte de données variées.

Mettre en œuvre des régressions de richesse croissante en matière de données explicatives conduit à évaluer l'amélioration de notre capacité à prédire des valeurs d'indicateurs de services écosystémiques. En effet, de manière générale, plus la régression a été réalisée à partir d'informations variées, plus le score de validation croisée tend à diminuer (augmentation de l'inverse du score – figure 2).

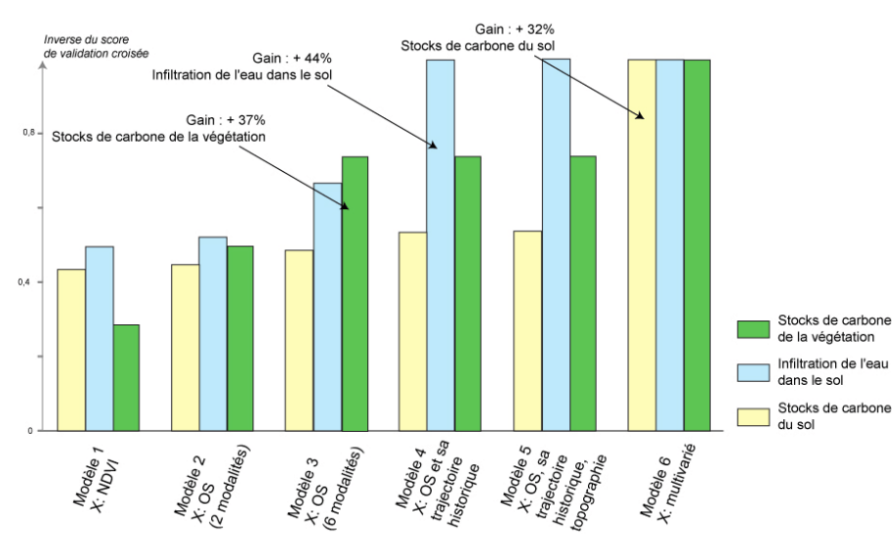


Figure 2 : Relation entre l'apport en données d'entrée dans la régression et augmentation des performances prédictives de cette dernière. Les scores de validation croisée ont été standardisés et ramenés à leur inverse, afin de faciliter la lecture.

Le NDVI, donnée la plus simple en matière d'obtention et de quantité d'informations, du moins en milieu tropical, est la moins performante pour expliquer les variations de ces indicateurs de services écosystémiques (table 5). Une occupation du sol en deux modalités, opposant forêts et

non forêts, est un peu plus compliquée à obtenir. Son utilisation engendre une forte diminution du score de validation croisée, particulièrement pour l'indicateur d'infiltration de l'eau dans le sol. Le passage à une régression dont la variable explicative est une classification du sol détaillée en six modalités engendre une rupture dans la tendance (diagramme en barres) de validation croisée, très nette pour l'indicateur des stocks de carbone de la végétation et beaucoup plus douce pour les deux autres indicateurs. Pour l'infiltration de l'eau dans le sol, la rupture est provoquée par l'ajout de l'information sur la trajectoire historique d'occupation du sol. Ce résultat met en avant le rôle, sur le moyen terme, de la couverture végétale ainsi que l'impact de la présence de bétail dans les propriétés biophysiques du sol. Pour chacun des trois indicateurs de services, la meilleure relation obtenue dans le cadre de cette étude est celle prenant en compte plusieurs variables. Pour les stocks de carbone de la végétation, le modèle le plus performant retient deux variables : le site et l'occupation du sol. Dans le cas de l'infiltration de l'eau dans le sol, le modèle le plus performant retient également deux variables : la trajectoire historique d'occupation du sol et l'occupation du sol en soi. Enfin, dans le cas des stocks de carbone du sol, le modèle le plus performant retient six variables : l'occupation du sol, la trajectoire historique d'occupation du sol, le site, le NDWI, la distance à l'eau et la topographie.

La présence de la variable « site » dans les régressions les plus performantes des stocks de carbone de la végétation et du sol montre qu'il existe un effet de contexte fort (effet site), essentiellement dû aux fortes différences géologiques et pédologiques entre Palmares et les deux autres localités. Les différences socio-économiques et politiques peuvent également expliquer cet effet de contexte.

| Variable(s)<br>explicative(s)                                                               | Carbone de la<br>végétation | Infiltration de l'eau<br>dans le sol    | Carbone du<br>sol                                                           |
|---------------------------------------------------------------------------------------------|-----------------------------|-----------------------------------------|-----------------------------------------------------------------------------|
| NDVI                                                                                        | NDVI                        |                                         |                                                                             |
| OS en 2 modalités                                                                           | OS                          |                                         |                                                                             |
| OS en 6 modalités                                                                           | OS                          |                                         |                                                                             |
| OS et Trajectoire<br>d'OS                                                                   | OS                          | Trajectoire d'OS et<br>OS (3 modalités) | OS, trajectoire<br>d'OS                                                     |
| OS, trajectoire d'OS<br>et topographie                                                      | OS                          | Trajectoire d'OS et<br>OS (3 modalités) | OS, trajectoire<br>d'OS                                                     |
| OS, trajectoire d'OS,<br>NDVI, distance à l'eau,<br>topographie, altitude,<br>pente et site | OS et Site                  | Trajectoire d'OS et<br>OS (3 modalités) | OS, trajectoire<br>d'OS, site,<br>NDWI, distance<br>à l'eau,<br>topographie |

Table 5 : Variables retenues par les méthodes de régression visant à expliquer les variations du stock de carbone de la végétation, l'infiltration de l'eau dans le sol et les stocks de

carbone du sol par différentes variables spatiales explicatives. Acronyme : OS = occupation du sol et VC = validation croisée. Les régressions 4 à 6 ont fait l'objet d'une sélection de variables, c'est-à-dire que seules les variables influençant significativement l'indicateur de services écosystémiques ont été retenues.

La cartographie fondée sur des données nombreuses et variées présente donc une incertitude plus faible, même dans le cas d'indicateurs de services écosystémiques résultant de processus biophysiques relativement simples (comme les stocks de carbone de la végétation). Cette analyse souligne ainsi la nécessité de prendre en compte plusieurs variables de natures différentes pour modéliser les indicateurs de services écosystémiques. Toutefois, l'importance de l'acquisition de données doit être considérée avec prudence et l'apport de données supplémentaires devrait être une considération systématique préalable à la cartographie des services écosystémiques. Il est essentiel de garder en tête le compromis entre la plus-value réelle de l'utilisation d'un grand nombre de données et les coûts ainsi que le temps liés à leur acquisition et traitement (Reyers et al. 2015). Il est parfois préférable d'avoir une cartographie un peu plus simple, fondée sur moins de données, mais une cartographie facile à interpréter et fondée sur des données fiables (Jacobs et al. 2015).

### **Une régression à visée explicative.**

L'arbre de décision constitue la méthode de régression la plus performante pour expliquer les variations des stocks de carbone du sol. Par conséquent, cette méthode a été utilisée pour réaliser une analyse complémentaire. Ce choix a également été influencé par la dimension pédagogique des arbres de décision, qui présentent l'intérêt d'être faciles à comprendre et à interpréter grâce à la hiérarchisation visuelle des variables qu'ils offrent. Dans le cas du carbone stocké dans le sol, l'ajout de variables exhaustives permet d'améliorer considérablement notre capacité à expliquer les variations de l'indicateur observées (le coefficient de détermination –  $R^2$  – passe de 0,52 à 0,62, soit une augmentation d'environ 16 % – figure 3). L'arbre à visée purement explicative présente de nombreuses variables, dont celles présentes dans l'arbre de décision utilisé dans le processus de cartographie (arbre dit « prédictif »). Toutefois, cet arbre à visée purement explicative présente aussi de nouvelles variables. Certaines, comme le type de sol, sont très significatives, c'est-à-dire qu'elles permettent d'expliquer en grande partie le phénomène étudié. D'autres variables sont certes significatives mais se retrouvent plus bas dans l'arbre, comme le nombre d'années depuis le premier brûlis et le nombre de brûlis. Cela veut dire que bien qu'importantes, elles impactent moins le service.

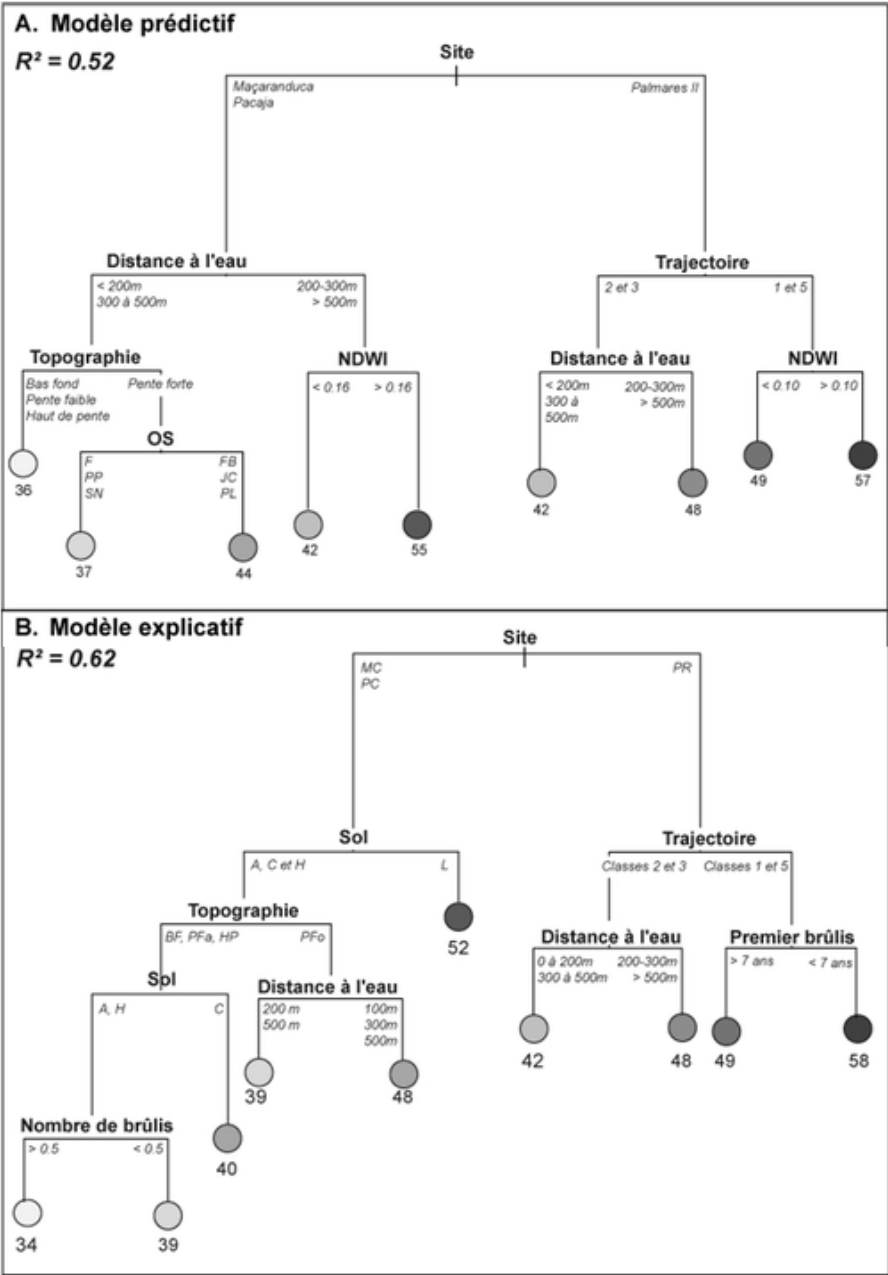


Figure 3 : A/ Arbre de décision prédictif des stocks de carbone du sol, réalisé à partir des données télédéteectées. B/ Arbre de décision à but non prédictif réalisé à partir des données télédéteectées et des données de terrain. Les variables « Sol » et « Premier brûlis » correspondent respectivement au type de sol et au nombre d'années depuis le premier brûlis.

## Discussion.

## De l'importance des choix méthodologiques.

La mise en évidence des choix méthodologiques qui sous-tendent l'exercice de cartographie d'indicateurs de services écosystémiques permet d'en souligner les potentialités et les limites. Les analyses proposées dans cet article sont fondées sur la mise en place de méthodes statistiques qui mettent en relation données de terrain (indicateurs de services écosystémiques) et données explicatives, essentiellement issues de la télédétection. Ces approches statistiques constituent une des démarches possibles de spatialisation des services écosystémiques. Les résultats de la modélisation statistique mettent en valeur plusieurs points. Tout d'abord, ils illustrent l'inégale capacité à spatialiser des indicateurs de services écosystémiques, à partir de données télédétektées et en se fondant sur les méthodes statistiques employées (Le Clec'h et al. 2016). Les cartographies d'indicateurs qui en résultent possèdent un taux d'incertitude lié à notre capacité à prédire des valeurs d'indicateurs par le biais d'implémentation de régressions, capacité qui est variable d'un indicateur à l'autre. En effet, les coefficients de détermination s'échelonnent de 0.52 à 0.75, montrant que certains indicateurs, tels que les stocks de carbone de la végétation, sont relativement faciles à modéliser à partir des données et des méthodes présentées, alors que d'autres, comme les stocks de carbone du sol, sont beaucoup plus difficiles à modéliser. Ensuite, les résultats statistiques ont montré la nécessité d'adapter la méthodologie à chaque indicateur de services écosystémiques. En effet, l'arbre de régression a été considéré comme la méthode la plus adaptée pour prédire les stocks de carbone du sol, alors que le modèle linéaire a été jugé plus pertinent pour prédire les stocks de carbone de la végétation et de l'infiltration de l'eau dans le sol. La capacité prédictive de chaque méthode employée est donc variable d'un indicateur à un autre et il n'est pas possible de déterminer *a priori* une méthode universellement meilleure. Par conséquent, il serait intéressant de tester d'autres types de méthodes statistiques et/ou d'approches – économiques (Busch et al. 2012) et sociales (van Riper et al. 2012).

Les analyses ont également illustré l'importance de ne pas fonder la modélisation d'indicateurs de services écosystémiques uniquement sur des données moins détaillées, même si celles-ci sont relativement faciles à obtenir (indice de végétation, classification d'occupation du sol). Ainsi la cartographie de certains indicateurs est grandement améliorée avec des informations plus difficiles et/ou plus longues à générer, comme celles liées à la dynamique de l'occupation du sol (par exemple, les stocks de carbone de la végétation), ou avec d'autres sources d'information comme celles liées à la topographie (par exemple, les stocks de carbone du sol pour lesquels la distance au réseau hydrographique est essentielle). L'ajout de données permet d'améliorer

significativement notre capacité à prédire des valeurs par application de méthodes statistiques, sans que cela se fasse de manière homogène d'un indicateur à l'autre. En effet, à partir de l'étude de trois indicateurs de services écosystémiques (stocks de carbone de la végétation et du sol et infiltration de l'eau dans le sol), nous avons montré que la qualité de la cartographie via l'application de méthodes statistiques prédictives est systématiquement sensible à la hausse de la qualité et/ou à l'augmentation du nombre de données. Ces dernières n'influencent pas de la même manière notre capacité à prédire de nouvelles valeurs pour les différents indicateurs. En effet, l'ajout de certaines variables, comme la trajectoire historique d'occupation du sol, peut ainsi s'avérer important pour certains indicateurs (infiltration de l'eau dans le sol) et inutile dans le cas d'autres (stocks de carbone de la végétation). Toutefois, à partir des données de télédétection utilisées dans cette étude, la capacité explicative maximale est rapidement atteinte et l'ajout de nouvelles données devient difficile et/ou n'apporte aucune amélioration significative. Bien que permettant d'obtenir des résultats satisfaisants, les données télédétectées montrent des limites dans leur capacité à expliquer les variations des indicateurs de services écosystémiques, notamment ceux liés aux processus biophysiques du sol, tels que les stocks de carbone du sol. Il est donc nécessaire de posséder un jeu de données riche et d'adapter la méthodologie à chaque indicateur de services écosystémiques. À l'instar de Chan et al. (2006), nous avons mis en évidence l'importance des méthodes multivariées, comprenant à la fois des données liées à la végétation et des données liées à l'hydrographie et à la topographie. Pourtant, les cartographies d'indicateurs de services écosystémiques se font sur des jeux de données moins variés, souvent centrés sur une seule thématique. Par exemple, dans la littérature scientifique, les études traitant de la question de l'infiltration de l'eau dans le sol et des stocks de carbone du sol se centrent sur des jeux de données peu variés en termes de thématiques, uniquement relatifs à la pédologie et/ou à la climatologie (Demarcin et al. 2011) (Maes et al. 2012). D'autres informations, comme celles obtenues auprès des acteurs du territoire, pourraient permettre d'apporter un autre point de vue à l'analyse. Elles offrent en effet l'opportunité de réaliser des cartographies participatives (Klain et Chan 2012) (Palomo et al. 2013) (Drakou et al. 2015) qui pourraient permettre (1) de posséder davantage de données, plus variées que celles uniquement accessibles par télédétection et (2) de faciliter la mise en place de discussions autour des services écosystémiques entre les différents acteurs du territoire (Paudyal et al. 2015). L'analyse a également mis en évidence l'importance de posséder des informations détaillées, notamment en termes d'occupation du sol (c'est-à-dire, une classification d'occupation du sol à plusieurs modalités, sans une opposition dichotomique urbain / milieu naturel ou forêt / non forêt). Or, une classification précise d'occupation du sol est plus difficile à réaliser à grande échelle (Domac 2004) (Foody 2015). Par conséquent, il semble plus facile de réaliser des cartographies

d'indicateurs de services écosystémiques à l'échelle locale. En outre, travailler à une échelle locale facilite l'obtention de données variées, telles que celles relatives au paysage, à la topographie et, dans certains cas, au climat ou aux grandes unités géologiques, ce qui permet d'obtenir des résultats statistiques plus précis. L'intérêt de travailler à une échelle locale a également été souligné par la présence de la variable « site » dans les régressions les plus performantes des stocks de carbone de la végétation et du sol. Ce fort effet de contexte reflète les différences en termes d'environnement biophysique, de politiques publiques ou encore d'historicité de la déforestation. Il met ainsi en avant l'importance des missions de terrain, permettant de mesurer localement les indicateurs afin d'obtenir des cartographies prenant en compte les spécificités locales. Ces missions de terrain sont d'autant plus importantes dans une perspective d'*empowerment* des acteurs locaux, via leur participation à l'élaboration d'une cartographie. Toutefois, comme souligné précédemment, la diversité et la précision des données utilisées, ainsi que la fiabilité de la cartographie, doivent systématiquement être pensées en relation avec le coût que l'acquisition de ces données représente.

### **Pour une critique constructive : « the hatchet and the seed ».**

Les difficultés à cartographier les services écosystémiques reflètent les flous tant théoriques que méthodologiques associés à cette notion. Il existe ainsi toute une série de choix (données, méthodologie ...) relatifs à l'exercice de modélisation et de cartographie des services écosystémiques. Ces choix sont à l'origine d'incertitudes. Ces incertitudes sont inhérentes à l'exercice, puisqu'un modèle et une carte sont des généralisations de l'information, des schématisations de la réalité. Or ces choix méthodologiques sont encore peu étudiés (Kandziora, Burkhard et Müller 2013) (Lavorel et al. 2014). Pourtant, le concept de services écosystémiques s'est progressivement imposé dans le domaine environnemental, chez les scientifiques mais aussi parmi les décideurs politiques ou les organismes internationaux. Proposer un regard critique sur ces choix méthodologiques est une première contribution à l'analyse critique de la notion même de services écosystémiques – « the hatchet » de Robbins (2004). Mais accompagner cette critique d'un regard constructif (« the seed ») permet aussi de relever l'intérêt de l'outil comme de la notion étudiés. Ainsi, considérant que la critique ne peut être réellement impactante que si elle est constructive et ne rabat pas la modélisation à un système opaque de boîte noire, il apparaît comme indispensable de développer des modélisations concurrentes qui permettent les comparaisons et les discussions de modèles (Bouleau 2014). Dès lors, la modélisation critique, en pénétrant les fondements des processus de fabrique des mises en cartes et en chiffres des objets de la Nature, peut alimenter la cartographie critique en

l'accompagnant dans son décryptage des contextes précis de production, d'analyse, de diffusion et d'usages des métrologies de l'espace, afin d'en révéler les enjeux socio-politiques.