

• Étude des risques de déforestation
dans et autour de la réserve de Luki

Modèle d'économétrie spatiale

Maxence Rageade
ONF International





TABLE DES MATIERES

TABLE DES MATIERES	2
TABLEAUX ET FIGURES	2
INTRODUCTION	3
METHODOLOGIE	4
Données.....	4
Échantillonnage.....	7
Modèle empirique.....	8
RESULTATS ET IMPLICATIONS.....	9
Résultats statistiques.....	9
Interprétation des résultats	12
Implications	13
CONCLUSION.....	15
REFERENCES.....	19

TABLEAUX ET FIGURES

Table 1: Prédiction globale du modèle versus observations de déforestation	9
Table 2 : Qualité du modèle de probabilité estimé	9
Table 3 : Résultats de l'estimation de la probabilité de déforestation par un modèle linéaire généralisé lié à une fonction logit.....	11
Figure 1 : Deforestation historique entre 2002 et 2012 (source OSFAC).....	16
Figure 2 : Simulation de la carte de risque de déforestation en 2012 pour l'ensemble de la région	17
Figure 3 : Carte de risque de déforestation simulée en 2012 pour la RB.....	18



INTRODUCTION

La zone d'étude comprend la réserve de biosphère de Luki (RBL) et les zones alentours en République Démocratique du Congo (RDC). La RBL a été créée dans l'objectif de concilier la conservation de la biodiversité et le développement des populations environnantes. Pour ce faire, ont été délimitées une zone centrale dédiée à la conservation et dans laquelle aucune coupe de bois n'est autorisée, une zone tampon dédiée à l'éducation à l'environnement, et une zone de transition où les activités humaines sont soumises à autorisation. Au-delà de la RBL, les activités ne sont plus soumises à restriction d'usage.

Historiquement, les dynamiques de déforestation sont importantes dans la région¹ (cf. Fig1). En 2002, l'ensemble de la région était couvert par environ 65 % de zones de forêts, alors qu'en 2012 on ne comptait plus que 50 % de couverture forestière. Ainsi sur la période, ce sont 52 500 hectares de forêt qui ont été convertis. Il existe cependant des différences notables entre les différentes zones composant la région. La zone centrale de la RBL ne comprend quasiment pas de déforestation, et la zone de transition a enregistré une perte de moins de 1 400 hectares de forêts. La grande majorité des zones déboisées se trouvent donc à l'extérieur de la RBL qui durant la période 2002-2012 a perdu plus de 30 % de son couvert forestier. Cette tendance laisse à craindre à l'avenir à un report des dynamiques de déforestation vers la RBL.

L'expansion agricole à travers les pratiques sur abattis-brûlis, et le prélèvement de bois dans la forêt pour la production de charbon sont les principales causes de la perte et la dégradation du couvert forestier dans la région. Si la croissance démographique, ainsi principale cause sous-jacente du changement d'usage des terres, ne ralentit pas, des actions ad hoc de lutte contre la déforestation (amélioration des pratiques agricoles, de l'efficacité énergétique de la consommation de bois de feu, etc.) doivent être entreprises pour espérer conserver la RBL. Néanmoins, une interrogation persiste quant à savoir où pourraient être précisément mises en place ces activités de lutte contre la déforestation.

Il est clair que pour lutter efficacement contre la déforestation, les activités doivent être mises en place en priorité sur les zones de forêts les plus à risque. Il convient donc, à partir des données empiriques, d'identifier les facteurs de risque de localisation de la déforestation et d'estimer leur impact sur la probabilité d'une forêt à être déboisée. Pour ce faire, cette étude propose un modèle statistique de localisation de la déforestation spécifique à la région.

Dans une première partie, nous préciserons la méthode et les données utilisées et dans une deuxième partie, nous décrirons les résultats obtenus et les enseignements qui peuvent en être tirés.

¹ L'analyse de la déforestation historique a été faite à partir des images Landsat interprétées par L'OSFAC pour produire des cartes de couvert forestier en 2002 et 2012.

METHODOLOGIE

Pour identifier les risques de localisation de la déforestation et ses facteurs, nous proposons dans cette étude d'estimer un modèle d'économétrie spatiale spécifique à la région de la RBL et ses alentours. Cette méthode empirique utilise les avancées statistiques les plus récentes et garantit la robustesse des résultats (Paelinck, 1979). Elle est donc particulièrement adaptée et utilisée pour comprendre les dynamiques de localisation de la déforestation et fournir des implications politiques (Ferretti-Galon, 2014). Néanmoins, la pertinence de ces méthodes d'estimations est largement dépendante de la disponibilité et la qualité des données utilisées. C'est pourquoi avant de présenter la méthode de régression utilisée pour paramétrer le modèle, nous décrivons dans la prochaine section les données disponibles et utilisées.

Données

Variable expliquée

Les cartes de Forêt Non-Forêt de 2002 et 2012, produites par l'OSFAC sur la base d'images Landsat dans le cadre du projet CBFF Luki, ont été utilisées pour créer la carte de déforestation sur cette même période. Cette carte de déforestation a été encodée pour créer une variable binaire (valeur 1 pour les zones déboisées et 0 sinon). Cette variable transformée (déforestation vs sans changement) est la variable expliquée de notre modèle.

Les cartes de Forêt Non-Forêt utilisées ont une résolution de 30 mètres. Dès lors, chaque observation de la variable déforestation a la même résolution. Or pris à une échelle locale (ici le pixel), la déforestation peut être caractérisée par un processus allotopique important c'est-à-dire qu'un changement ou non d'usage des sols localisé dans un espace donné est souvent en partie expliqué par des facteurs causaux localisés dans d'autres espaces voisins. En d'autres termes la déforestation enregistrée à l'échelle d'un pixel sur la carte à toutes les chances d'être expliqué par la dynamique de déforestation de ses pixels voisins. Ce phénomène peut avoir plusieurs explications. Il peut exister un effet de propagation de déforestation sur une observation lié à la déforestation des observations voisines, notamment si la surface moyenne des patchs de déforestation dans la région est supérieure à la résolution de l'observation, ou si les techniques d'abattis-brûlis sont mal maîtrisées (propagation non maîtrisée du feu de forêt). Dans tous les cas, les changements observés sur les zones voisines sont un signe d'une pression anthropique qui se rapproche et donc d'un risque de déforestation plus important. Il est donc essentiel de prendre en compte ce phénomène, non seulement pour ces considérations causales, mais aussi d'un point de vue statistique.

Car en effet, les modèles d'économétrie spatiale stipulent le principe d'interdépendance spatiale (Le Gallo, 2002). Dès lors nous devons nous assurer pour éviter des biais d'estimation qu'il n'existe pas d'autocorrélation spatiale, ou dans le cas contraire que celle-ci soit prise en compte. Dans notre cas, après avoir effectué le test de Moran, la présence d'autocorrélation spatiale n'a pu être rejetée (Moran, 1948). Nous utilisons donc dans le modèle la variable de déforestation, pour créer une variable endogène décalée.

Variable endogène décalée

Afin de prendre en considération l'autocorrélation spatiale dans le modèle, nous utilisons un modèle autorégressif spatial en incluant une variable endogène décalée. Puisque les observations spatiales de déforestation sont admises être interdépendantes, nous considérons dans le modèle leur position relative les unes par rapport aux autres. Ainsi nous spécifions la topologie du système spatial en construisant une matrice de poids général exogène à partir d'une fonction de distance euclidienne inverse entre les observations et bornée à 2000 mètres² (Bavaud, 1998).

Avec d_{ij} la distance entre l'observation i et j , on obtient la matrice de poids à partir de :

$$w_{ij} = \begin{cases} \frac{1}{d_{ij}} & \text{si } d_{ij} < 2000m \\ 0 & \text{sinon} \end{cases}$$

Nous relierons ensuite cette matrice de poids à la variable dépendante de déforestation y_j pour former la variable endogène décalée de l'observation i , comme suit :

$$[Wy]_i = \sum_{j=1}^N w_{ij} y_j$$

Cet opérateur spatial est introduit comme variable explicative dans l'estimation économétrique, ce qui permet de corriger les biais liés à la présence d'autocorrélation spatiale.

Variables explicatives

Les variables explicatives utilisées dans le modèle sont classées selon leur relation causale à la déforestation usuellement admise (Kaimowitz, 1998). Ainsi dans la région de la RBL nous avons segmenté les facteurs de localisation de la déforestation de la manière suivante :

- Les facteurs de contraintes zonales
- Les facteurs d'accessibilité
- Les facteurs d'activités anthropiques
- Les facteurs biophysiques

Les facteurs de contraintes zonales représentent des unités administratives clairement identifiées et qui sont susceptibles de favoriser ou non la déforestation. Ces facteurs sont introduits dans le modèle à travers des variables muettes qui prennent la valeur de 1 lorsque l'observation est à l'intérieur de la zone et 0 sinon.

Les premières unités géographiques retenues sont les différentes zones de la RBL : la zone centrale, la zone tampon et la zone de transition. Les enclaves – zones d'habitation au sein de la zone de transition à l'intérieur desquelles les activités agricoles sont autorisées - au sein de la zone de transition seront également différenciées. Ces distinctions nous permettront

² Pour des questions de faisabilité nous limitons la taille de la matrice en introduisant une distance de voisinage maximale. Ainsi nous admettons que la déforestation d'une observation n'a pas d'impact sur les observations situées au delà de 2km.

d'identifier si ces zones contraintes ont un impact différent sur le risque de déforestation et d'estimer ainsi leur efficacité respective.

Une troisième zone a été introduite dans le modèle, c'est **la concession forestière**³ (nommée Agirumbe dans la région) située au nord-ouest de la région d'étude. L'effet de cette concession est à priori incertain dans la mesure où celle-ci pourrait être un frein à la déforestation car d'accès limité comparée aux zones sans contrainte (a priori soumise à un plan d'aménagement) (Megevand, 2013). Mais cette concession pourrait au contraire ne pas respecter les principes d'un prélèvement soutenable ou encore attirer la convoitise des populations locales pour le prélèvement de bois illégal. Dans ce cas, la zone sous concession serait un facteur de déforestation. Actuellement l'exploitation dans la concession se limite à la collecte de noix de palme pour la fabrication d'huile. Les activités de gestion sont donc limitées.

Les facteurs d'accessibilité sont l'ensemble des réseaux infrastructureux ou hydrologiques qui ouvrent l'accès aux zones forestières et qui sont a priori des facteurs importants de localisation de la déforestation (Scouvard, 2006). Ce peut être aussi des voies d'échange et d'accès aux marchés qui diminuent les coûts d'acheminements et accroissent ainsi l'incitation à la déforestation des forêts proches.

Les réseaux routiers⁴ sont, dans la majorité des études sur la déforestation, identifiés comme, un facteur majeur de déforestation. Dans la région, il existe trois types de réseaux routiers : les routes locales, provinciales et nationales. Afin de distinguer leur impact respectif, nous avons différencié ces trois types de routes dans le modèle. Par ailleurs, nous ne cherchons pas à mesurer l'impact direct d'une construction d'une route sur le changement d'usage des sols (c.-à-d. l'empreinte de la route) mais l'impact de la proximité d'une route sur la décision de déforestation des populations locales. Pour ce faire, pour chacun des trois types de routes nous construisons une variable représentant la distance (euclidienne) à la route la plus proche pour chacune des observations de forêts. La distance à la route peut avoir une relation non linéaire avec la probabilité de déforestation. Ainsi, nous pouvons a priori supposer qu'à partir d'une certaine distance de la route, il n'y a plus d'impact ou au contraire que l'impact sur la probabilité de déforestation redevient positif. Car en effet, l'enclavement de certaines zones pourrait expliquer l'établissement de campements et ainsi des activités de déforestation loin des routes. Pour prendre en considération ce potentiel effet, nous introduisons dans la régression la distance aux routes au carré.

De manière identique, à travers le calcul de la distance euclidienne, nous construisons une variable de proximité **aux réseaux hydrologiques**⁵. Alors que les réseaux hydrologiques peuvent être, comme pour les routes, un facteur d'accessibilité, et d'échange dans le cas de voies navigables, ils ont la caractéristique additionnelle d'être un proxy de la proximité des zones humides (ou irrigables) qui pourraient avoir une influence sur le développement agricole : défavorable aux environs très proches à cause de l'inondation annuelle temporaire des terres et effet favorable pour l'implantation des villages et/ou campements.

Certaines variables révèlent la présence d'une **activité humaine** continue qui menace la préservation des forêts à proximité. Ce sont par exemple, les zones habitées et les marchés économiques qui sont généralement admis comme étant une cause majeure de déforestation (Angelsen, 2007). Plus particulièrement, la proximité à ces villes/villages contribuerait à diminuer le coût d'approvisionnement des productions locales (cf. approche Von Thünen

³ Sources des données cartographiques: WRI 2013

⁴ Sources des données cartographiques: WRI 2013

⁵ Sources des données cartographiques: WRI 2013

appliquée à la déforestation). Dès lors, pour cette étude, *la distance aux villes/villages*⁶ les plus proches est incluse dans le modèle. Par ailleurs, nous savons que la principale place d'échange dans la région est la ville de Matadi et que celle-ci peut donc avoir un impact différent de l'impact moyen de l'ensemble des autres localités. C'est pourquoi nous prenons en compte dans le modèle une variable représentative *de la distance à la ville de Matadi (et donc aussi à Kinshasa)*.

Alors que les villes/villages sont une zone d'écoulement des productions agricoles, il est essentiel d'identifier les zones où ces dernières sont produites et ainsi estimer où leur développement à venir aura lieu. Car, le développement agricole, principale cause de déforestation dans la région, a lieu le plus souvent à proximité des zones agricoles déjà installées. C'est pourquoi disposant d'une carte d'utilisation des terres historique dans laquelle sont identifiées les zones de développement agricole en 2002, nous produisons une variable de *proximité aux zones déjà cultivées*⁷. Calculée comme la distance euclidienne des observations forêts à la zone en culture la plus proche, cette variable devrait nous permettre d'identifier les zones alentours les plus propices à l'expansion agricole à venir et ainsi à la déforestation.

Un dernier facteur, signe de la proximité de la présence d'activités humaines en cours (production de charbon de bois ou de bois de chauffe par exemple) ou passées, est *la distance aux sols nus*, même si ce facteur pourrait tout aussi bien être classé dans les facteurs d'accessibilité. Car en effet, les zones de forêts à proximité des zones déjà ouvertes (c.-à-d. à proximité des sols nus) sont de fait plus accessibles et donc présentent a priori un risque plus important d'être déboisées que les zones de forêts intérieures. Nous prenons donc en compte dans le modèle une variable de distance euclidienne aux sols nus, convertie en logarithme népérien pour des raisons techniques.

Enfin, nous admettons que l'état des facteurs biophysiques peut influencer la décision de conversion des forêts (Geist, 2001). Ainsi deux facteurs biophysiques sont pris en compte par le modèle comme probable facteur de déforestation : *l'altitude*⁸ et *les pentes*. L'altitude peut être un facteur de localisation des activités de déforestation en favorisant l'implantation des cultures et/ou des zones d'habitation sur des altitudes relativement faibles. Même si la région ne présente pas une forte variabilité topographique, nous introduisons l'altitude dans le modèle comme variable de contrôle. Par ailleurs, les pentes (créées à partir du modèle d'altitude) sont aussi prises en compte dans le modèle. On s'attend ainsi que le risque de déforestation soit plus limité sur les zones accidentées ou au contraire que les zones relativement plates soient propices à l'installation des surfaces agricoles et/ou des campements et favorisent ainsi le changement d'usage des sols.

Échantillonnage

D'un point de vue technique, il n'est pas possible d'estimer le modèle économétrique sur l'ensemble des observations (c.-à-d. des pixels) de la région⁹. Nous avons donc procédé à un échantillonnage des observations de forêt en 2002 dans la région. Pour ce faire nous avons

⁶ Sources des données cartographiques: WRI 2013

⁷ Sources des données cartographiques: OSFAC

⁸ Sources des données cartographiques: NASA / CGIAR CSI

⁹ L'ensemble des observations classées comme forêt représentent environ 2.8 millions de pixels. L'ensemble de ces observations ne peut pas être traité avec des logiciels statistiques standard et dans des délais raisonnables notamment en raison de la difficulté pour produire la matrice spatiale de poids.

construit un échantillon systématique de 181 811 observations de forêt en 2002. Ceci correspond à un écart de 100 mètres entre les observations ce qui est suffisamment petit pour avoir une matrice de pondération spatiale fiable et suffisamment grande pour satisfaire les considérations statistiques.

Modèle empirique

Nous avons utilisé un Modèle Linéaire Généralisé lié à une fonction Logit pour estimer la probabilité de déforestation d'une observation en fonction des variables décrites précédemment (Fahrmeir, 1994). La variable dépendante est la déforestation ($y_i = 1$ si l'observation a été déboisée, $y_i = 0$ sinon) sur la période 2002-2012. Toutes les observations de non-forêt en 2002 ont été exclues de l'échantillon de manière à s'assurer que $y_i = 0$ ne correspond qu'aux zones de forêts qui sont restées forêts sur la période et non aux sols nus qui sont restés sols nus.

Nous modélisons la déforestation pour chaque observation (i) à travers une loi binomiale de Bernoulli avec une probabilité (p_i) d'être déboisée en fonction de la valeur des variables exogènes associées à (i). Par ailleurs, après avoir décelé la présence d'autocorrélation spatiale, nous décidons d'introduire la variable endogène décalée comme expliquée précédemment.

En somme, la forme réduite du modèle estimé peut s'écrire:

$$\ln\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta_k x_{k,i} + \rho W_i y_i + \varepsilon_i$$

avec α la constante du modèle, β_k les paramètres des k variables explicatives x_k ; ρ est le paramètre de la variable endogène décalée $W_i y_i$ de l'observation i ; et ε le terme d'erreur.



RESULTATS ET IMPLICATIONS

Résultats statistiques

Les résultats obtenus pour estimer la probabilité de déforestation dans la région de RBL sont globalement significatifs à un seuil d'incertitude inférieur à 1%. Par ailleurs, le pouvoir prédictif du modèle ici mesuré par le ROC (Receiver Operator Curve) est de 92.9 % (Grey, 1972). Le tableau résume la qualité des prédictions du modèle de probabilité en comparant la déforestation estimée (probabilité > 0.5) versus la déforestation réellement observée.

Table 1: Prédiction globale du modèle versus observations de déforestation

		Observation		Total
		Déboisée (D)	Non Déboisée (ND)	
Prédiction	Déboisée (+)	32 062	7 650	39 712
	Non Déboisée (-)	15 046	127 053	142 099
Total		47 108	134 703	181 811

Comme le confirme le tableau 2¹⁰ le modèle a une légère tendance à sous-estimer la probabilité de déforestation des observations qui ont réellement été déboisées avec néanmoins un taux de sensibilité de plus de 68%. Le modèle n'a au contraire aucune difficulté à identifier les zones avec peu de risque de déforestation et qui n'ont effectivement pas été déboisées. Dans ce cas la spécificité est de plus de 94%. Notons qu'en somme les prédictions correctes du modèle estimé sont de plus de 87%.

Table 2 : Qualité du modèle de probabilité estimé

		Taux
Sensibilité	$Pr(+ D)$	68%
Spécificité	$Pr(- ND)$	94.3%
Valeur prédictive positive	$Pr(D +)$	80.7%
Valeur prédictive négative	$Pr(ND -)$	89.4%
Taux de faux + pour vrai ND	$Pr(+ ND)$	5.7%
Taux de faux - rate pour vrai D	$Pr(- D)$	31.9%
Taux de faux + pour prédit +	$Pr(ND +)$	19.3%
Taux de faux - pour prédit -	$Pr(D -)$	10.6%
Classifié correctement		87.5%

Le tableau 3 reprend l'ensemble des paramètres estimés pour les différentes variables explicatives de la probabilité de déforestation dans la région de la RBL. La seconde colonne du

¹⁰ La seconde colonne du tableau 2 explique la définition des indicateurs de la colonne 1. Avec (Pr) la probabilité, (+) et (-) le nombre d'observations de forêts prédit comme déboisé et non déboisé respectivement, et (D) et (ND) le nombre d'observations effectivement observé déboisée et non déboisée respectivement. Tous ces nombres (+), (-), (D) et (ND) sont tirés du tableau 1.

tableau correspond aux coefficients estimés de chacune des variables de la colonne 1. Les colonnes suivantes reprennent les tests statistiques habituels appliqués à chacune des variables expliquées de la régression économétrique (Erreur standard, le test z, la probabilité de non significativité ($P > |Z|$), et l'intervalle de confiance à un seuil de 95%). De manière générale plus le $P > |Z|$ est proche de 0 plus la variable explicative associée est significative. Ainsi, la majorité des variables sont significatifs à un seuil de 1%. Seule la variable, « zone transition » n'est pas significative à un seuil acceptable ($>25\%$).

Plusieurs combinaisons de variables ont été testées pour évaluer la robustesse des résultats. Le modèle, ici présenté, a été sélectionné à partir des informations de l'AIC (Akaike Information Criterion) et du BIC (Bayesian Information Criterion) de Schwarz.

Table 3 : Résultats de l'estimation de la probabilité de déforestation par un modèle linéaire généralisé lié à une fonction logit

<i>Pr (Déforestation)</i>	Coefficient	Err. Std	<i>z</i>	<i>P> z </i>	[95% interval Conf.]
<i>Spatial lag</i>					
Zone centrale RBL	6255.281	61.03811	102.48	0.000	[6135.648 6374.913]
Zone transition RBL	-1.843579	.2606624	-7.07	0.000	[-2.354468 -1.33269]
Zone tampon RBL	-.0460479	.0410322	-1.12	0.262	[-.1264695 .0343737]
Zones enclaves	-.7980952	.1305079	-6.12	0.000	[-1.053886 -.5423045]
Titre forestier	.1851399	.0850503	2.18	0.029	[.0184443 .3518355]
Pentes	-.695671	.0436759	-15.93	0.000	[-.7812741 -.6100679]
Dist hydrologique	-.0031251	.0009744	-3.21	0.001	[-.005035 -.0012153]
Dist localite	-5.37e-06	2.07e-06	-2.60	0.009	[-9.42e-06 -1.32e-06]
Dist Matadi	-.0000109	4.55e-06	-2.39	0.017	[-.0000198 -1.97e-06]
Dist routes locales	-.0000145	1.01e-06	-14.46	0.000	[-.0000165 -.0000126]
Dist routes provinciales	-.0000443	6.04e-06	-7.33	0.000	[-.0000561 -.0000324]
Dist routes nationales	-.0000297	4.34e-06	-6.85	0.000	[-.0000382 -.0000212]
Altitude	-.0000279	6.18e-06	-4.50	0.000	[-.00004 -.0000157]
Dist zones agricoles	.0012633	.0000901	14.03	0.000	[.0010868 .0014398]
Dist routes locales (squared)	-.0004203	.0000199	-21.08	0.000	[-.0004594 -.0003812]
Dist routes provinciales (squared)	1.88e-09	3.70e-10	5.09	0.000	[1.15e-09 2.61e-09]
Dist routes nationales (squared)	6.93e-10	1.55e-10	4.46	0.000	[3.89e-10 9.97e-10]
Dist sols nus (Ln)	1.28e-09	3.79e-10	3.39	0.001	[5.41e-10 2.03e-09]
Constante	-1.56014	.0146634	-106.40	0.000	[-1.588879 -1.5314]
	4.874173	.1114848	43.72	0.000	[4.655667 5.092679]

Interprétation des résultats

L'importance des zones contraintes

Nos résultats montrent clairement que la zone centrale et la zone de transition de la RBL ont toutes deux un impact négatif sur la probabilité de déforestation (cf. Figure 3). En d'autres termes entre 2002 et 2012 ces deux zones ont joué efficacement leur rôle de contrainte sur le changement d'utilisation des terres. À noter néanmoins que la zone centrale de la RBL a un impact sur la conservation des forêts beaucoup plus important que la zone de transition. En terme d'impact marginal¹¹, la présence d'une zone de forêt dans la zone centrale de la RBL diminue le risque de déforestation de plus de 10% ce qui est considérable. La zone tampon a un impact plus mesuré mais non négligeable car elle réduit la probabilité de déforestation de plus de 6%. Au contraire les zones d'enclaves augmentent en moyenne le risque de déforestation de plus de 2%.

La zone sous concession forestière au nord-ouest de la RBL joue un rôle significatif de conservation des forêts supérieur à la zone de transition. Ceci s'explique certainement par le contrôle sur l'accès aux terres que le concessionnaire a encore même si l'activité d'exploitation est actuellement réduite. L'impact marginal sur la réduction de la probabilité de déforestation est d'environ 5%.

Le rôle essentiel des dynamiques voisines

Parmi les variables impactant le plus la probabilité de déforestation, on note la variable spatiale endogène décalée et la proximité des forêts aux zones ouvertes (c.-à-d. aux sols nus). Le modèle montre sans équivoque que si une zone de forêt est située à proximité d'une zone de forêt en cours de déboisement, sa probabilité de déforestation augmente à son tour considérablement.

À noter de la même manière que les zones de forêt ouvertes présentent un risque de déforestation beaucoup plus élevé que les zones de forêts intérieures. Ainsi selon l'impact marginal estimé, c.-à-d. si la distance d'une frontière de forêt aux zones de non forêt diminue de 1% la zone de forêt concernée voit sa probabilité de déforestation accroître de 15%.

Attention à l'expansion agricole

Sans surprise, la proximité des forêts aux surfaces agricoles (identifiées sur les cartes de couvert de l'OSFAC) augmente le risque de déforestation. Ce résultat trivial confirme bien le risque que représente l'expansion des surfaces agricoles sur les forêts alentour.

L'impact substantiel des voies de communications et des places d'échanges

L'ensemble des facteurs d'accessibilité introduit dans le modèle est positif et significatif. En d'autres termes, la proximité des réseaux routiers et hydrologiques augmente, significativement et substantiellement, la probabilité de déforestation. Cependant, le modèle met en avant des différences notoires en termes d'impact entre les différentes variables. Ainsi

¹¹ L'impact marginal représente la sensibilité de la variable expliquée à une modification à la marge de la variable explicative. Ces impacts marginaux ont été estimés de manière ad-hoc pour certaines variables uniquement notamment en raison des difficultés d'interprétations pour les autres et particulièrement pour les variables de distance.

la proximité aux routes locales est la variable la plus décisive sur la décision de déforestation. Cet impact est suivi néanmoins de près par l'effet de la proximité aux routes provinciales et nationales qui ont toutes deux un impact identique. À noter par ailleurs que la distance aux routes a un effet non linéaire. En effet, si les forêts à proximité des routes sont en péril, la réciproque n'est pas forcément vraie. C'est un résultat intéressant qui traduit qu'à partir d'une certaine distance des routes, la probabilité de déforestation augmente. Cet effet peut expliquer la déforestation sur des zones de forêts enclavées, loin des réseaux routiers et qui peut inciter les populations locales à mettre en place des campements le temps d'une campagne de culture par exemple.

La proximité au réseau hydrologique est elle aussi significative même si son impact sur le choix de localisation de la déforestation est moins sensible qu'à la proximité du réseau routier.

Les résultats de l'estimation confirment tout aussi bien la pression que présente la proximité des localités dans la région. À noter par ailleurs que Matadi, la principale place d'échange économique dans la région, a un impact sur la probabilité de déforestation qui est supérieur à l'impact moyen enregistré pour l'ensemble des autres localités.

Les facteurs biophysiques significatifs mais faiblement impactant

Les forêts situées sur des zones accidentées semblent être relativement préservées du risque de déforestation. Ceci confirme les difficultés de mise en place d'activités anthropiques sur des zones pentues.

Plus surprenant l'altitude a un impact positif sur la probabilité de déforestation. Ainsi plus l'altitude est élevée plus le risque de déforestation est important. Cependant l'impact marginal faible et la faible variabilité de l'altitude dans la zone nécessitent néanmoins d'interpréter ce résultat avec prudence. Il semble en effet que les zones les plus élevées situées à l'est de la zone d'étude concordent avec les zones où la déforestation est la plus importante. Néanmoins dans ce cas la causalité n'est pas établie.

L'intensité de ces facteurs sur la probabilité de déforestation reste néanmoins relativement faible.

Implications

Afin de discuter des enseignements et potentielles implications que peut offrir le modèle, nous avons produit à partir des résultats économétriques une carte de risque de déforestation simulée¹² sur les forêts de 2012.

De manière globale, la carte de risque de déforestation laisse apparaître un front de déforestation (ou arc de déforestation) partant du Nord-Ouest, passant par le Sud et remontant au Nord-Est. Ainsi même si la zone centrale de la RBL reste pour l'instant relativement épargnée, on peut anticiper une avancée de ce front dans les années à venir (d'autant plus rapide que la dynamique de déforestation sera importante), menaçant ainsi la zone de préservation de la RBL.

¹² Cette carte de risque a été produite à partir de la combinaison des cartes des différents facteurs et les résultats des paramètres économétriques. Le modèle de probabilité de déforestation a été simulé sur Dinamica Ego.

Une analyse plus fine de la carte de risque de déforestation simulée à partir des résultats économétriques nous permet de formuler trois points d'attention majeurs sur les dynamiques de déforestation à venir, notamment pour la zone centrale de la RBL.

Les forêts morcelées en péril

Il apparaît clairement sur la carte (fig. 2) que les risques de déforestation sont significativement plus élevés aux endroits où la forêt est déjà morcelée. Ces mosaïques de forêts à haut risque, surtout présentes au Sud et à l'Est de la région d'étude, traduisent principalement l'impact des facteurs de propagation de la déforestation et de l'ouverture des forêts en front des zones déjà déboisées. Une vigilance et un suivi spécifique sur ces mosaïques de forêts ouvertes sont recommandés. Il semble néanmoins difficile à ce stade d'avancement de pouvoir éviter leur déforestation, à moins de parvenir à réduire drastiquement la quantité de déforestation dans la région, car ce devrait être les premières forêts à être déboisées à l'avenir.

La zone centrale sous protection (menacée) de la zone de transition

La zone centrale est relativement bien préservée du risque de déforestation imminent (fig. 3). Néanmoins, la conservation de cette zone est largement dépendante de la dynamique de déforestation dans la zone de transition. Or celle-ci semble plus que menacée particulièrement au Sud et le long de la route nationale qui traverse la zone du Sud-Ouest au Nord-Est. Si les forêts du Sud de la zone centrale venaient à s'ouvrir (c.-à-d. si la déforestation dans la zone de transition se poursuit), la seule contrainte de préservation offerte par la zone centrale ne pourrait limiter le risque de déforestation, comme ce fût le cas jusqu'à présent. Dans l'immédiat, il semble opportun de mettre en place un plan de prévention et de suivi de la forêt et des activités humaines à l'intérieur d'une zone de buffer autour du tronçon de route qui traverse la zone de transition.

Forêt mitée, forêt en danger

Le mitage à l'intérieur des zones de forêt conduit à des risques de déforestation à l'intérieur de ces zones. Ces mitages sont particulièrement néfastes pour les forêts, car elles sont généralement difficilement observables sur le terrain et peuvent conduire à former des galeries qui peuvent à terme mosaïquer les zones de forêt et ainsi augmenter considérablement le risque de déforestation (cf. discussion sur les forêts morcelées). Ce risque de déforestation est par exemple très présent à la limite Est de la zone centrale de la RBL où le mitage est important. L'on peut aussi deviner certaine galerie en cours de formation, par exemple à la limite Ouest de la zone centrale.

CONCLUSION

Cette étude sur les risques de déforestation a été établie à partir d'un modèle d'économétrie spatiale sur la probabilité de déforestation. Les résultats sont significatifs et robustes. Ils ont ainsi permis de produire une carte de risque de déforestation fiable à partir de laquelle il est possible d'identifier les zones à risques et de deviner les tendances à venir. Ainsi si la zone centrale de la RBL est aujourd'hui relativement hors danger, la zone de transition devrait être soumise à une déforestation accrue dans les années à venir. Sans intervention ad hoc, le risque de déforestation devrait à moyen terme se propager dans la zone centrale au fur et à mesure que la zone de transition se découvre. Il est donc fortement recommandé d'entreprendre des actions supplémentaires dans la zone de transition qui reste pour l'instant une barrière de protection de la zone centrale.

Cette étude se limite à estimer les risques de localisation de la déforestation et n'apporte pas d'élément sur les tendances à l'évolution de la quantité de déforestation. Si les résultats de cette étude permettent d'identifier les zones sur lesquelles les activités devraient être menées en priorité, il sera mené en complément une étude spécifique sur les causes et facteurs de la déforestation pour quantifier les tendances à venir et ainsi le type d'activité à mettre en place.

Figure 1 : Déforestation historique entre 2002 et 2012 (source OSFAC)

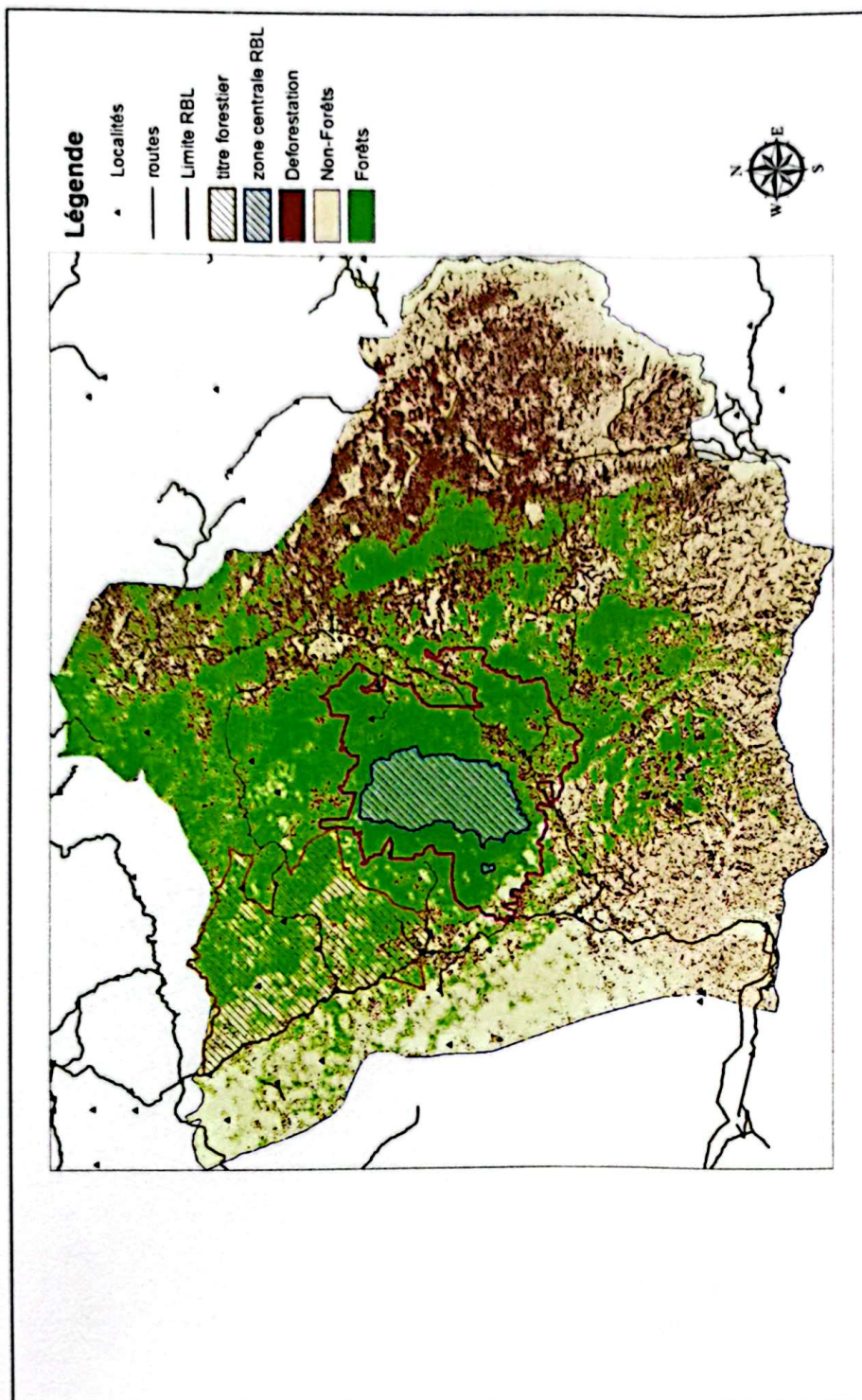


Figure 2 Simulation de la carte de risque de déforestation en 2012 pour l'ensemble de la région

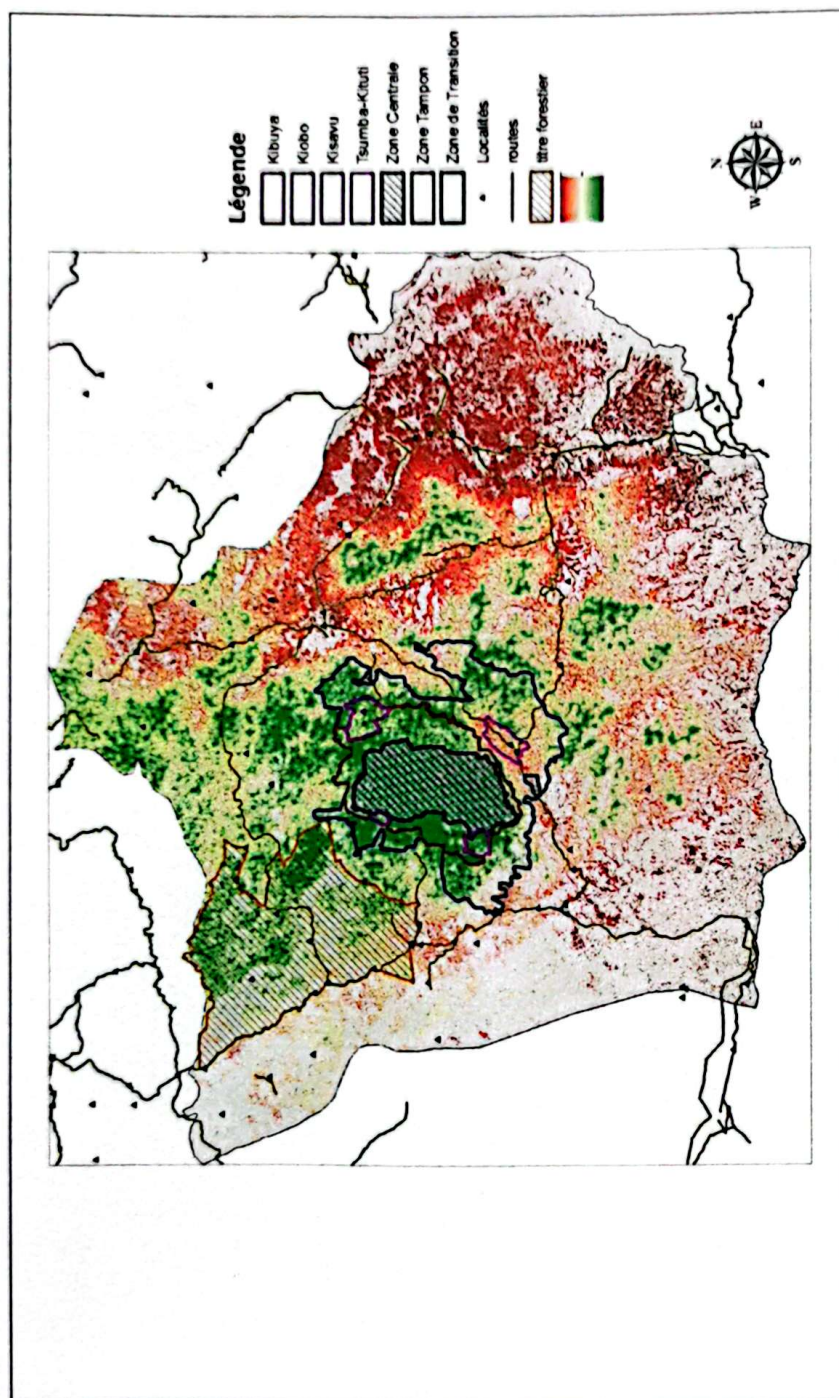
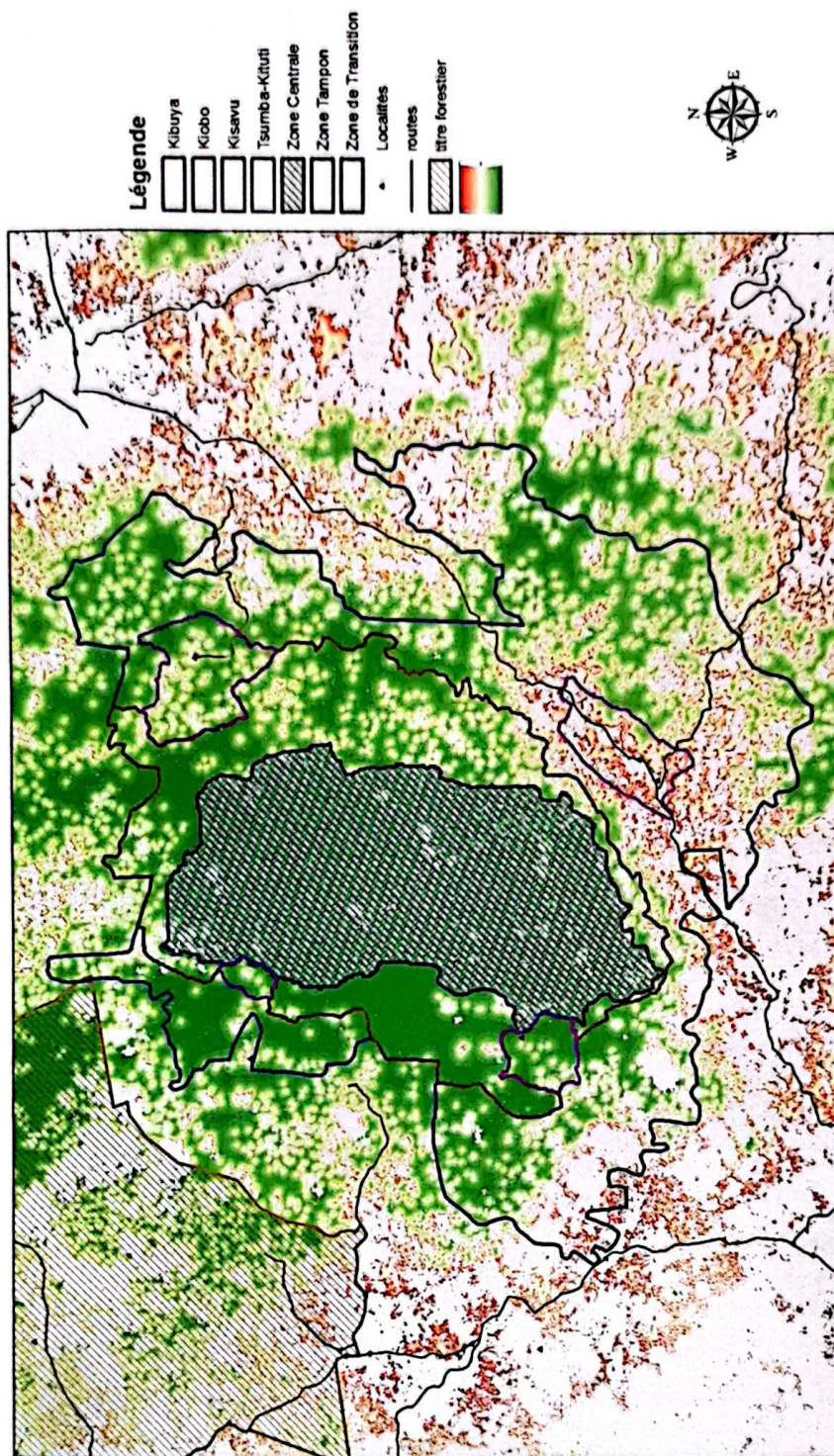


Figure 3 : Carte de risque de déforestation simulée en 2012 pour la RB



REFERENCES

- Angelsen, A. (2007). Forest cover change in space and time: combining the von Thünen and forest transition theories (Vol. 4117). World Bank Publications.
- Bavaud, F., 1998. Models for spatial weights: a systematic look, *Geographical Analysis*, 30, 153-171
- Fahrmeir, L., & Tutz, G. (1994). Multivariate statistical modelling based on generalized linear models (Vol. 2). New York: Springer.
- Ferretti-Galon, K., Busch, J., 2014. What Drives Deforestation and What stops it ? A Meta-Analysis of Spatially Explicit Econometric Studies. CGD Working Paper 361. Washington, DC : Center for Global Development
- Geist, H. J., & Lambin, E. F. (2001). What drives tropical deforestation. *LUCC Report series*, 4, 116.
- Grey, D. R., & Morgan, B. J. T. (1972). Some aspects of ROC curve-fitting: Normal and logistic models. *Journal of Mathematical Psychology*, 9(1), 128-139.
- Kaimowitz, D., & Angelsen, A. (1998). Economic models of tropical deforestation: a review. Cifor.
- Le Gallo, J. (2002). Econométrie spatiale: l'autocorrélation spatiale dans les modèles de régression linéaire. *Economie & prévision*, (4), 139-157.
- Megevand, C. (2013). Dynamiques de déforestation dans le bassin du Congo: Réconcilier la croissance économique et la protection de la forêt. World Bank Publications.
- Moran, P., 1948, the interpretation of statistical maps, *Journal of the Royal Statistical Society*, 10B, 243-251
- Nguyen Van, P., & Azomahou, T. (2003). Déforestation, croissance économique et population. *Revue économique*, 54(4), 835-855.
- ONFI, 2013. Validation par un standard carbone du projet pilote REDD+ autour de la réserve de biosphère de Luki en RDC – Etude de faisabilité.
- Paelinck, J.H.P, Klaassen, L.H., 1979. Spatial econometrics, Saxon House, Farnborough
- Scouart, M., & Lambin, É. F. (2006). Approche systémique des causes de la déforestation en Amazonie brésilienne: syndromes, synergies et rétroactions. *L'Espace géographique*, (3), 241-254.