



Universiteit
Leiden
The Netherlands

SMALL Savannah : an information system for the integrated analysis of land use change in the Far North of Cameroon

Fotsing, E.

Citation

Fotsing, E. (2009, December 8). *SMALL Savannah : an information system for the integrated analysis of land use change in the Far North of Cameroon*. Retrieved from <https://hdl.handle.net/1887/14619>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/14619>

Note: To cite this publication please use the final published version (if applicable).



Photo 9. Un groupe de femmes autour d'un point d'eau potable dans le village de Gadas, Extrême Nord du Cameroun.

Chapitre 9. Analyse spatiale et quantitative des facteurs déterminant les structures et changements d'utilisation de l'espace

Résumé

La connaissance des relations quantitatives entre la structure de l'utilisation de l'espace et les facteurs déterminants d'ordre biophysique et socioéconomique est une étape très importante dans la construction des modèles spatiaux dynamiques de type empirique. Le choix des variables à inclure dans ces modèles est le plus souvent difficile du fait de la complexité des processus en jeu dans les systèmes d'utilisation de l'espace. Ce chapitre décrit la mise en œuvre d'une méthode combinant les approches inductive et déductive pour identifier les facteurs les plus pertinents qui expliquent la structure et les changements d'utilisation de l'espace dans une région autour de Maroua dans l'Extrême Nord du Cameroun. A partir des couches d'information extraites de la base de données du SIE SMALL Savannah, l'analyse spatiale a permis de calculer des variables représentant au mieux les processus étudiés. La procédure de sélection des variables s'appuie sur une analyse de type régression logistique. Un modèle de prédiction de la structure et des changements au cours de la période 1987 et 1999 a été construit pour chacune des 6 utilisations de l'espace considérées dans la zone d'étude. Les résultats confirment la très forte influence du centre urbain et révèlent celle des potentialités des terres agricoles sur les formes d'utilisation de l'espace. La pression sur l'espace évaluée à partir d'un indice de population potentielle explique mieux les structures et les changements que la densité de la population. Cette application fournit des éléments de compréhension du système d'utilisation de l'espace de la région et les informations quantitatives obtenues sont utilisées pour construire un modèle qui simule les dynamiques spatiales liées aux différentes trajectoires d'évolution.

Mots clés : structure d'utilisation de l'espace, facteur déterminant, approche déductive/inductive, régression logistique, modèle spatial empirique

Abstract

The knowledge of quantitative relationships between land use patterns and biophysical as well as socioeconomic driving factors is an important step in building dynamic GIS-based models. More often, the choice of variables to include in these models is difficult due to the complexity of the system and the underlying processes. This chapter describes the implementation of a method combining inductive and deductive approaches to identify the most pertinent sets of factors that explain land use patterns and change in a region around Maroua, Far North of Cameroon. Land use and driving factors data were derived from the SMALL Savannah EIS database. Spatial analysis is used to calculate variables that better represent the land use processes. The procedure of variable selection is based on a logistic regression model. A predictive model of land use pattern and change between 1987 and 1999 was built for each of the 6 land use types considered in the region. Results confirm the strong influence of the main urban centre and reveal that of soil suitability on land use types. Population pressure evaluated as a population potential index, explains land use patterns and changes more adequately than population density. This study has provides key elements for understanding the land use system of the study area. Quantitative information derived from this application is used to build a model that simulates spatial dynamics related to different trajectories of land use change.

Key words: land use pattern, driving factor, deductive/inductive approach, scale, logistic regression, empirical GIS-based model

9.1. Introduction

L'exploration spatiale explicite et l'analyse quantitative des relations entre l'utilisation de l'espace et ses facteurs déterminants, sont très importantes dans la mesure où elles fournissent des éléments de compréhension des processus en jeu dans le système agraire. De plus, les informations quantitatives sur la manière dont les facteurs déterminants interagissent avec l'utilisation de l'espace sont requises pour supporter la modélisation dynamique des changements futurs de l'utilisation de l'espace et de leur effet. Pendant longtemps, les approches utilisées pour étudier ces questions ont été le plus souvent orientées par les traditions disciplinaires. Les études basées sur de telles approches focalisent l'attention sur un nombre limité de facteurs et fournissent une explication partielle des phénomènes. De plus en plus, les études adoptent des approches pluridisciplinaires qui se focalisent sur l'explication des processus sans se soucier des disciplines. Dans ce cas, plusieurs points de vue sont intégrés dans l'analyse qui prend en compte la multiplicité des facteurs déterminants pour construire des modèles prédictifs plus pertinents des changements d'utilisation de l'espace (Veldkamp et Fresco, 1996).

Cette thèse propose une telle approche intégrée pour analyser les changements d'utilisation de l'espace en zone des savanes de l'Extrême Nord du Cameroun. Aux chapitres 5 et 7, une analyse préalable a permis de caractériser les grandes dynamiques agraires et de produire une liste de facteurs qui peuvent potentiellement expliquer les changements observés. Ce travail préalable qui combine les informations issues de la revue de la littérature avec les données secondaires et les observations de terrain est important pour la compréhension des dynamiques en cours mais, elle ne fournit pas d'informations précises et localisées. Le chapitre précédent a abordé la résolution de ce problème en permettant l'identification et la quantification des principales conversions de l'occupation du sol. Toutefois, l'analyse des logiques d'organisation des formes d'occupation du sol et des causes des changements observées est restée descriptive, qualitative et hypothétique. Les explications étant principalement basées sur l'interprétation visuelle des images, les observations de terrain et des discussions avec les acteurs impliqués dans l'utilisation et la gestion de l'espace. Une telle approche ne permet pas toujours de capturer toutes les interactions des facteurs et les chaînes de causalité qui gouvernent les structures d'utilisation de l'espace et les changements observés sur une vaste région. On court le plus souvent le risque de généraliser un phénomène local à l'échelle de la région.

Le présent chapitre propose une approche d'analyse complémentaire pour mieux comprendre les facteurs qui déterminent la structure et les dynamiques du système d'utilisation de l'espace. La méthode vise à identifier les combinaisons les plus pertinentes des facteurs explicatifs de la structure et des changements de l'utilisation de l'espace. Une quantification de leur influence devrait permettre de construire des modèles prédictifs performants de la distribution spatiale des changements d'utilisation de l'espace. La performance des modèles, fait ici référence aussi bien à la précision de localisation, la rationalité et l'interprétabilité des résultats dans le contexte de la zone d'étude. Pour atteindre ces objectifs, la méthode d'exploration suivie combine les aspects d'une approche de type inductif avec les éléments d'un raisonnement de type déductif. Le contenu du chapitre est organisé de la façon suivante : la section suivante rappelle le contexte et le cadre méthodologique de l'étude ; puisque celle-ci repose sur le cadre conceptuel et les outils proposés par le Système d'Information SMALL Savannah, les sections suivantes correspondent plus ou moins aux résultats de l'application de chaque module de SMALL Savannah à cette étude de cas. La section 3 décrit l'étape de préparation des données et de représentation des phénomènes

étudiés. La section 4 décrit les résultats de la mise en œuvre de la procédure de sélection des facteurs déterminants sur le jeu de données issu de SMALL Savannah. La section 5 propose une synthèse et une discussion des résultats. La section 6 dégage les perspectives de recherches futures.

9.2. Contexte et cadre méthodologique de l'étude

9.2.1. Zone d'étude et complexité des changements d'utilisation de l'espace

La zone d'étude choisie pour cette application du SIE SMALL Savannah est la même que celle qui a été décrite au chapitre précédent. Une spécificité de cette zone est l'unicité du centre urbain de Maroua, situé au cœur de cette région de l'Extrême Nord du Cameroun. Cette ville aurait une très grande influence sur l'utilisation de l'espace autorisant la formulation d'une hypothèse d'organisation de l'espace autour du centre urbain tel qu'énoncé dans la théorie de l'état isolé de Von Thünen (1826). D'une superficie d'environ 3600 km², la population totale de cette petite région est estimée à 500 000 habitants en 2000, soit une densité moyenne de 138 hab./km² (RGPH, 1987). En dehors du centre urbain qui compte environ 200 000 habitants (Seignobos et al., 2001), c'est une zone de densité moyenne comparée à la zone des monts Mandara voisine située à l'ouest où on enregistre une très forte densité de population. On distingue trois principaux groupes ethniques dans cette petite région. Les Foulbé représente 29%, les Guiziga représentent 21%, les Mofu 34% et les autres ethnies représentent 16%. Nous avons montré au chapitre 5 que chaque ethnie avait une conception différente de l'organisation et de la gestion de l'espace qui pourrait influencer les dynamiques observées. Les mouvements migratoires les plus récents observés au cours des deux dernières décennies parfois stimulés par la mise en place d'infrastructure de communications, traduisent essentiellement des réponses à l'augmentation de la pression démographique, à la dégradation locale des ressources naturelles et aux nouveaux besoins économiques des populations rurales. Ces migrations récentes ont des conséquences sur les changements d'occupation du sol et d'utilisation de l'espace et peuvent conduire à d'importantes mutations inattendues dans le système agraire. On a remarqué par exemple que les migrations engendrent dans les zones d'accueil un processus d'extension des terres cultivées qui se traduit dans le paysage par un défrichement important des savanes boisées. Ces pratiques agricoles extensives imposées par la pression démographique et la faible disponibilité des terres exposent les sols à la dégradation. Les terroirs où aboutissent les courants migratoires les plus importants connaissent ainsi très vite une situation de saturation foncière qui peut conduire à la dégradation ou à des formes plus intensives d'utilisation de l'espace (Milleville et Serpantié, 1994).

Au chapitre 5, une analyse préalable a permis de caractériser les grandes dynamiques agraires et de produire une liste de facteurs qui peuvent potentiellement expliquer les changements observés. Ce travail préalable qui combine les informations issues de la revue de la littérature avec les données secondaires et les observations de terrain est important pour la compréhension des dynamiques en cours mais, elle ne fournit pas d'informations précises et localisées. Le chapitre précédent a abordé la résolution de ce problème en permettant l'identification et la quantification des principales conversions de l'occupation du sol. Toutefois, l'analyse des logiques d'organisation des formes d'occupation du sol et des causes des changements observés est restée descriptive, qualitative et hypothétique. Les explications étant principalement basées sur l'interprétation visuelle des images, les observations de terrain et des

discussions avec les acteurs impliqués dans l'utilisation et la gestion de l'espace. Une telle approche ne permet pas toujours de capturer toutes les interactions des facteurs et les chaînes de causalité qui gouvernent les structures d'utilisation de l'espace et les changements observés sur une vaste région. On court le plus souvent le risque de généraliser un phénomène local à l'échelle de la région.

Le présent chapitre porte sur la mise en œuvre d'une approche d'analyse complémentaire pour mieux comprendre les facteurs qui déterminent la structure et les dynamiques du système d'utilisation de l'espace. Cette application vise à identifier les combinaisons les plus pertinentes des facteurs explicatifs de la structure et des changements de l'utilisation de l'espace. Une quantification de leur influence devrait permettre de construire des modèles prédictifs performants de la distribution spatiale des changements d'utilisation de l'espace. La performance des modèles faits ici référence aussi bien à la précision de localisation, la rationalité et l'interprétabilité des résultats dans le contexte de la zone d'étude.

9.2.2. Justification de l'approche d'analyse des changements d'utilisation de l'espace

Les changements d'utilisation de l'espace sont le plus souvent la résultante d'une combinaison complexe d'un nombre important de facteurs qui changent dans le temps et en fonction des échelles et niveaux d'organisation. L'un des préalables d'une meilleure interprétation des processus observés dans les systèmes agraires est la reconnaissance de cette complexité et la mise en œuvre d'approches d'analyse appropriées. La communauté des chercheurs travaillant sur les changements d'utilisation de l'espace a pris conscience de cet enjeu. Il est maintenant reconnu que la structure et les changements d'utilisation de l'espace sont déterminés dans le temps et l'espace par une combinaison de facteurs de nature biophysique, sociodémographique, géoéconomique et politique qui interagissent à différentes échelles spatiales (Turner et al., 1995). Ces facteurs peuvent influencer directement l'utilisation de l'espace où à travers une longue chaîne de causalité qui peuvent également opérer à différents niveaux d'organisation. Les questions qui se posent dans différents contextes géographiques concernent l'identification des facteurs, l'évaluation de leur importance relative et les mécanismes qui sont en jeu à différentes échelles.

Au chapitre 2 de cette thèse nous avons montré que les approches utilisées pour aborder ces questions ont été le plus souvent orientées par les traditions disciplinaires. De plus en plus, les études adoptent des approches pluridisciplinaires qui se focalisent sur l'explication des processus sans se soucier des disciplines. Dans ce cas, plusieurs points de vue sont intégrés dans l'analyse qui prend en compte la multiplicité des facteurs déterminants pour construire des modèles explicatifs et prédictifs des changements d'utilisation de l'espace (Veldkamp et Fresco, 1996). Le choix des variables à considérer dans la construction de ce type de modèles est le plus souvent difficile du fait de la complexité du système et des processus à modéliser. Les approches scientifiques utilisées pour aborder ce problème peuvent être classées dans un continuum qui se situe entre deux extrêmes (Overmars et al., 2006). On distingue d'un côté, les approches purement inductives qui correspondent aux situations où le choix des facteurs ne s'appuie ni sur un modèle ou une théorie quelconque. La méthode consiste dans ce cas à explorer les grandes bases de données à la recherche de corrélations qui permettraient de déduire des explications. Un premier pas de cette extrémité vers les approches déductives consiste à combiner l'exploration des données avec des hypothèses sur les facteurs qui pourraient être pertinents à l'explication des processus étudiés. Un deuxième pas consiste à définir les variables explicatives en s'appuyant sur des théories de changement d'utilisation de l'espace existantes. Un troisième pas consiste à définir également la structure du modèle en

s'appuyant sur la théorie qui a permis d'identifier les variables. D'autres approches intermédiaires peuvent être définies en incluant un niveau plus ou moins important d'éléments d'une approche déductive. A l'autre extrémité, les approches purement deductives sont celles où pour un système d'utilisation de l'espace donnée, un modèle ou une théorie est spécifié et les paramètres définis. Le modèle ou la théorie est ensuite testé par rapport à la réalité afin d'en apporter des améliorations.

La plupart des études qui portent sur la compréhension des changements d'utilisation de l'espace ont une forte tendance inductive alors que très peu s'appuient sur des théories existantes pour identifier et quantifier les facteurs explicatifs. Les approches inductives sont certes appropriées dans les situations où on ne dispose pas ou de très peu de connaissances sur le système étudié mais, un inconvénient majeur c'est qu'elles peuvent conduire à de nombreuses corrélations qui ne correspondent pas nécessairement à des mécanismes ou à des processus pertinents. L'inconvénient des approches deductives c'est qu'elles réduisent le plus souvent la réalité et arrive difficilement à saisir toute la complexité des systèmes étudiés. Toutefois, elles présentent l'avantage d'incorporer explicitement dans les modèles les mécanismes en jeu, de fournir des meilleures preuves de causalité et de produire des structures qui peuvent être facilement généralisées. En somme, une approche plus réaliste devrait consister à combiner les deux approches en tirant profit des avantages tout en minimisant les inconvénients de chacune.

9.2.3. Cadre méthodologique de l'étude

Pour atteindre les objectifs énoncés dans ce chapitre, la méthode d'exploration suivie dans cette application combine les aspects d'une approche de type inductif avec les éléments d'un raisonnement de type deductif. En effet, la méthode combine l'exploration systématique des données avec les théories de changement d'utilisation de l'espace (chapitre 2) et les connaissances empiriques sur le système d'utilisation de l'espace de la zone d'étude (chapitre 5). Ce chapitre illustre ainsi un exemple d'application du SIE SMALL Savannah présenté au chapitre 6. La figure 9.1 montre que les principaux modules du SIE mis en œuvre sont le module d'analyse spatiale et le module d'explication et prédiction. Le diagnostic du système d'utilisation de l'espace qui a été effectué au chapitre 5 n'est plus traité ici. Les paragraphes suivant décrivent les étapes de la procédure d'identification des facteurs déterminants et de construction des modèles de prédiction.

L'objectif de cette procédure est de sélectionner pour chaque utilisation de l'espace, l'ensemble le plus pertinent de facteurs qui expliquent la structure et les changements observés. Dans plusieurs études, le choix des facteurs déterminants les structures d'utilisation de l'espace et leur changement est limité par la disponibilité des données. On suppose le plus souvent que les facteurs déterminants la localisation des utilisations de l'espace à un moment donné, seront celles qui vont opérer pour une période donnée. L'inconvénient de cette situation est que les prédictions au-delà de cette période ne sont pas valides. Pour cette application, l'analyse statistique est effectuée sur deux dates (1987 et 1999). Ceci permet dans un premier temps de prendre en compte le changement de la nature et de l'importance des facteurs dans l'explication de la structure spatiale de l'utilisation de l'espace. L'analyse statistique effectuée pour le changement entre les deux dates permet également d'expliquer la dynamique spatiale des changements observés. L'idée principale de la méthode d'exploration est de tirer profit à la fois des connaissances empiriques sur les processus et des possibilités offertes par les données spatiales explicites disponibles.

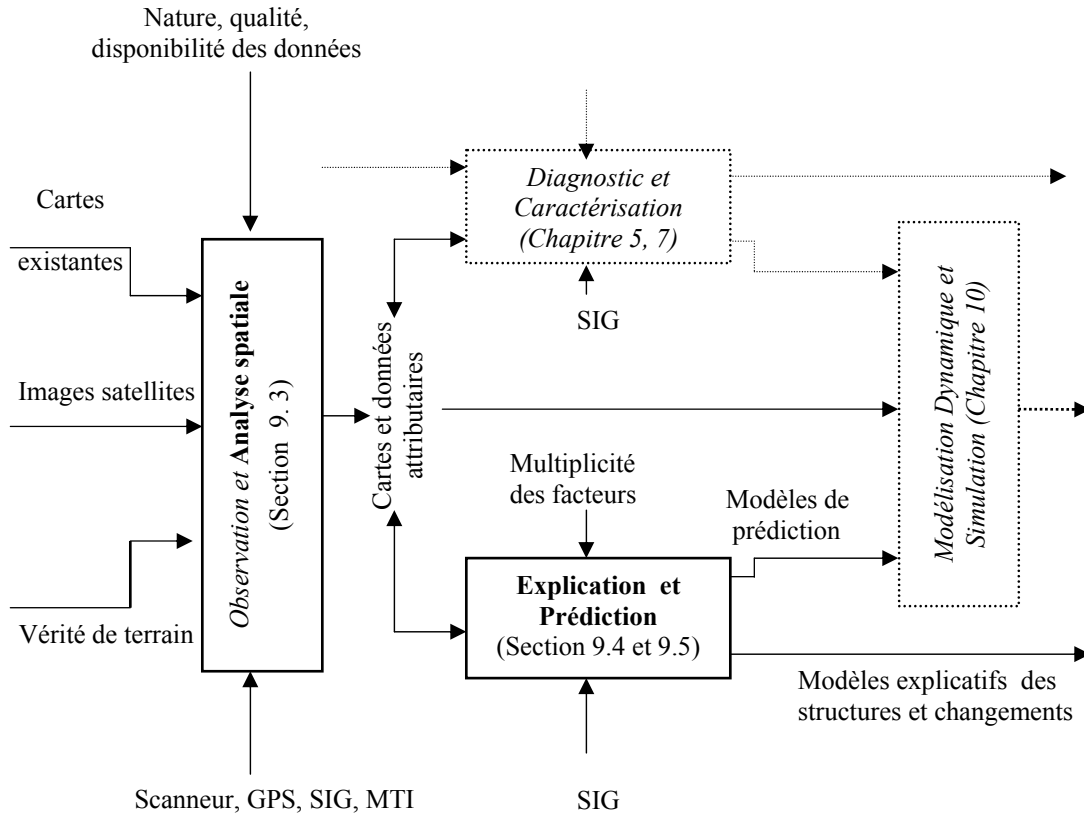


Figure 9.1 : Cadre conceptuel et outils de SMALL Savannah utilisés pour l'analyse des facteurs déterminants les changements d'utilisation de l'espace. Les rectangles en pointillé et les textes en italique représentent les modules non utilisés dans cette application.

Les principales étapes de l'identification des facteurs déterminants incluent :

- la construction des régressions simples entre l'utilisation de l'espace et toutes les variables;
- la sélection des variables potentiellement déterminantes ;
- l'exécution de la procédure de régression conditionnelle multiple pas à pas ;
- la construction et l'évaluation des performances du modèle de prédiction final.

Interprétation des paramètres statistiques des régressions simples

Si on considère l'équation du modèle logistique introduit au chapitre 3, les principaux paramètres statistiques utilisés pour la sélection des variables sont : les coefficients de corrélations ou coefficients logistiques (β_i), la signification statistique de la corrélation (P) et le pouvoir explicatif de la variable dépendante mesuré par la méthode ROC. Le signe et la valeur des coefficients β_i déterminent la nature et l'importance de la contribution du facteur représenté par la variable X_i . En effet pour des coefficients β négatifs, lorsque la valeur de la variable indépendante X grandit, $\beta_0 + X\beta$ devient grand et la probabilité de présence de l'utilisation de l'espace tend vers 1. Pour des coefficients positifs, elle tend vers 0 lorsque la valeur de la variable indépendante X_i grandit (Agresti et Finlay, 1997). Les variables de type catégorie

ayant n valeurs sont traitées comme n-1 variables binaires. Une valeur correspondant à une catégorie moyenne étant choisie comme référence dans la procédure du calcul de la régression. Dans les cas des variables *soilpot* et *soilsorg*, il s'agit respectivement de *soilpot4* et *soilsorg4*. Les coefficients logistiques sont donnés pour chacune des n-1 variables. Le test de Wald est utilisé pour apprécier la signification de chaque corrélation. Les corrélations sont jugées significatives pour $P < 0.01$, soit un degré de confiance de 99% (Agresti et Finlay, 1997). Dans cette application, les niveaux de signification des variables dérivées des variables de type catégorique ne sont pas considérés individuellement dans la procédure de sélection. La valeur de la superficie au-dessous de la courbe ROC est interprétée comme le pouvoir explicatif de la variable indépendante par rapport à la variable dépendante pour chaque modèle de régression simple (Pontius et Schneider, 2001).

Sélection des variables candidates et construction des modèles de prédiction

La procédure de sélection des facteurs déterminant potentiellement a été introduite au chapitre 3. On considère deux seuils de pouvoir explicatif minimum $ROC_{s1} = 0,6$ et $ROC_{s2} = 0,7$.

Les variables dépendantes qui se sont révélées au cours de l'exploration des données sont sélectionnées si le pouvoir explicatif est supérieur à 0,7. Toutefois, la corrélation devra correspondre à une relation pertinente lorsqu'on se réfère aux théories ou aux connaissances empiriques sur le système d'utilisation de l'espace dans la région. Les variables ayant un pouvoir explicatif compris entre 0,6 et 0,7 sont sélectionnées si elles sont déterminantes par hypothèse. Cela signifie que leur corrélation traduit un processus pertinent ou correspond à une hypothèse à tester. Les variables ainsi sélectionnées sont insérées dans une régression conditionnelle pas à pas de type ascendant. Cette procédure de régression évalue la contribution des différentes variables à l'explication de la variable dépendante et décide de l'insertion ou de l'exclusion d'une variable en s'appuyant sur le niveau de signification statistique et l'importance de la colinéarité avec les autres variables du modèle. On arrive ainsi à réduire les effets de multi-colinéarité et les variables indépendantes incluses sont faiblement corrélées les unes aux autres. Les variables finalement sélectionnées à l'issue de la régression pas à pas sont utilisées pour calculer l'équation du modèle de régression. Dans le cas de la régression multiple finale, la valeur en dessous de la courbe ROC, calculée pour chaque utilisation de l'espace en 1987 ou en 1999, est interprétée comme la capacité de discrimination du modèle de prédiction défini par l'équation de régression obtenue. Pour la carte des changements entre 1987 et 1999, elle est interprétée comme la capacité de détection des changements de cette utilisation de l'espace (Pontius et Schneider, 2001).

9.3. Représentation de l'utilisation de l'espace et des facteurs déterminants

Cette section décrit les traitements mis en œuvre pour représenter au mieux les phénomènes liés à l'utilisation de l'espace et aux facteurs déterminants. On distingue les traitements préalables des données et les opérations d'analyse spatiale

9.3.1. Traitements préalables des données

Les traitements préalables incluent l'extraction des couches d'information de la base de données du SIE SMALL Savannah et la définition d'une échelle commune de représentation des données (étendue et

résolution). Ces données étant disponibles à différentes échelles spatiales (chapitre 4), le traitement préalable vise également à obtenir un ensemble de données géoréférencées sur la zone d'étude.

Les limites de cette zone étant artificielles parce que déterminées par l'étendue des scènes d'image satellite utilisées, les effets de bordures sont très importants. A titre d'exemple, pour le calcul de la distance au marché ou au centre urbain le plus proche, les zones de bordures qui sont proches des marchés ou des centres urbains situés hors de la zone n'ont pas des valeurs réalistes. Pour réduire ces effets de bordure, le traitement de toutes les données c'est à dire le calcul des variables spatiales est effectué dans ces cas à l'échelle de la grande région avant d'extraire ensuite l'étendue correspondant à la zone d'étude.

Toutes les couches d'informations sont ramenées à la même résolution spatiale pour que les croisements soient possibles et l'interprétation pertinente. Le choix d'une résolution d'analyse doit être un compromis entre l'échelle d'opération des processus et la résolution des données disponibles sur l'utilisation de l'espace et les facteurs déterminants. Le choix de l'échelle d'analyse est difficile dans la mesure où tous les processus n'opèrent pas à la même échelle et chaque processus opère sur une plage d'échelle et non sur une échelle spécifique (Allen et Starr, 1982 ; Bian, 1997). Les images satellites utilisées ont une résolution de 20 m mais les cartes d'occupation du sol produites ont une résolution de 100 m qui se justifie par le niveau de précision de la méthode de cartographie (voir chapitre 8). A cette résolution, la carte contient encore des hétérogénéités ou détails qui sont difficiles à lier avec les facteurs déterminants d'autant plus que les couches d'information correspondantes ne sont pas disponibles à une échelle aussi fine.

Compte tenu de l'échelle de représentation des autres couches d'information (relief, type de sol, population), la résolution minimale a été fixée à 250 m. On a procédé au changement de résolution de cette carte en procédant à plusieurs agrégations spatiales successives avec un pas de 250 m. Les proportions de chaque classe d'utilisation de l'espace sont maintenues mais les structures d'utilisation de l'espace changent. A chaque pas de temps, les structures de la carte obtenue sont comparées avec celles de la carte de résolution initiale. La procédure de comparaison utilisée est celle proposée par Costanza (1989) et qui compare les structures entre deux images à plusieurs résolutions. L'algorithme de comparaison et de calcul du coefficient d'ajustement F_t a été introduit et décrit au chapitre 3. La taille des fenêtres de comparaison varie de 1 à 100 cellules avec un pas de 4. Le coefficient d'ajustement global (F_i) est calculé comme une moyenne pondérée des coefficients d'ajustement successifs de chaque fenêtre (F_w). Les courbes décrivant les mesures de similarité successives sont présentées à la figure 9.2. On remarque qu'au-delà de 1 000 m de résolution, l'ajustement global devient très bas ($F_t < 0.7$) et la courbe prend de plus en plus l'allure d'une droite indiquant que, les structures des deux images sont très différentes. Cette différence est d'ailleurs perceptible visuellement. Entre 250 m et 1000 m par contre on a des courbes traduisant une bonne similarité de structure entre les deux images. En effet, la courbe augmente d'abord très rapidement avant de se stabiliser à une valeur proche de 1. La plage de 250 à 100 m peut donc être utilisée en garantissant la conservation des grandes structures de l'utilisation de l'espace. Dans cette application, les données ont été ramenées à la résolution de 250 m. L'espace rectangulaire à analyser est donc une grille de 271 lignes et 288 colonnes, soit 78 048 cellules.

Un test de multi-colinéarité a été appliqué à toutes les variables incluses dans les modèles en effectuant une régression entre chaque variable indépendante et les autres variables indépendantes. Les données spatiales sur l'utilisation de l'espace ont une tendance à être dépendantes : on parle d'auto corrélation spatiale. L'autocorrection spatiale est la propriété qui fait que, du fait de la proximité géographique, des

variables dans l'espace (sur une certaine distance) prennent des valeurs plus proches qu'on ne l'aurait espéré si les observations étaient distribuées de façon aléatoire (Overmars et al., 2003). L'ensemble des analyses s'est effectué sur un échantillon aléatoire de 20% (soit 11380) des cellules avec les données de 1987 et 1999 afin de minimiser les effets de l'auto corrélation spatiale. Avec les données des changements d'utilisation de l'espace entre 1987 et 1999, l'analyse est faite sur toutes les cellules (57144) puisque le nombre total d'observation est très faible.

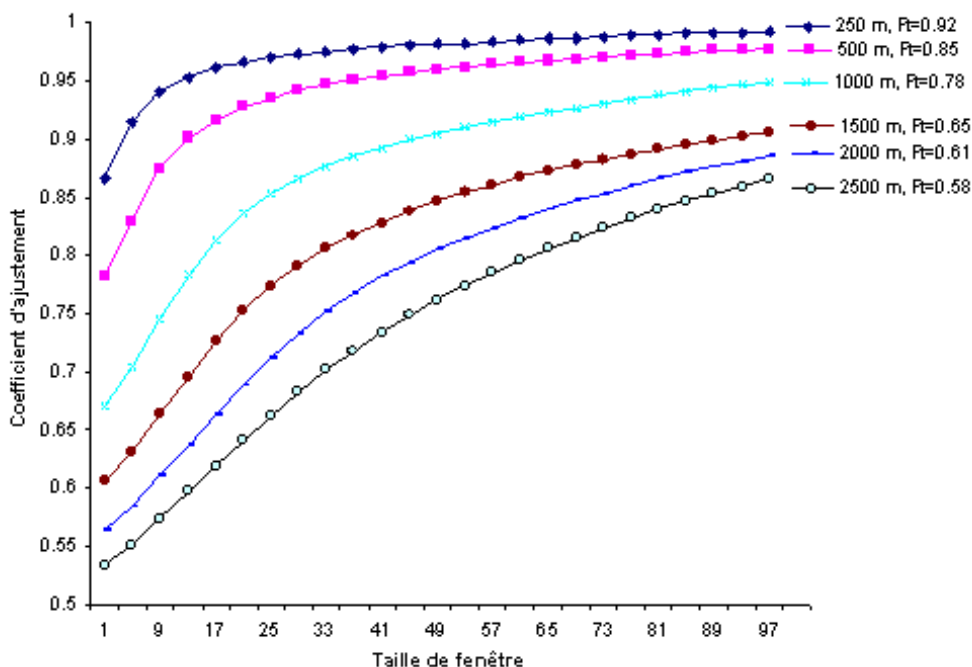


Figure 9.2 : Effets du changement de résolution spatiale sur les structures d'utilisation de l'espace. Les différentes cartes sont comparées avec la carte initiale de résolution 100 m.

9.3.2. Procédure de calcul des variables spatiales

Cette section, décrit comment les variables représentant l'utilisation de l'espace et les facteurs potentiellement déterminants, ont été formulées et représentées. La sélection des données sur les facteurs déterminants s'est faite sur la base d'une revue de la littérature. Une liste de facteurs qui peuvent expliquer les changements d'utilisations de l'espace a été ainsi produite au chapitre 5. Sur la base des connaissances disponibles sur la zone d'étude, une description est donnée sur la façon dont chaque facteur influence l'organisation ou les changements d'utilisation de l'espace. Lorsque des connaissances ne sont pas disponibles, des hypothèses sont émises sur les processus dominants. Ces hypothèses peuvent être dérivées de la théorie ou des expériences dans des zones similaires. De même, les variables ont été construites à partir des données en s'appuyant sur quelques théories de changement d'utilisation de l'espace qui ont été présentées au chapitre 2. Il s'agit par exemple de la théorie de Boserup (1965) qui aborde les conséquences de la densité de la population, la théorie de Malthus au sujet de l'impact de la population sur la dégradation environnementale (Hommer Dixon, 1999), la théorie de Von Thünen (1826) qui parle des marchés et leur influence sur l'organisation des utilisations de l'espace (O'Kelly et Bryan, 1996). En appliquant une suite de fonctions d'analyse spatiale, les variables quantitatives représentant au mieux les processus étudiés, sont calculées à partir des couches d'information issues de la base de données de SMALL Savannah. Les

opérations d'analyse spatiale ont permis de dériver un ensemble de 47 variables réparties de la façon suivante. 18 variables portent sur l'utilisation de l'espace et ses changements soit 6 types d'utilisation de l'espace pour chacune des trois situations : 1987, 1999 et les changements entre les deux dates. On dénombre 15 variables représentant les facteurs biophysiques, 8 variables représentant les facteurs sociodémographiques et 6 variables représentant les facteurs géoéconomiques. Les paragraphes suivant décrivent tour à tour la nature et la procédure de calcul de chacune de ces variables.

Utilisation de l'espace

Les variables représentant l'utilisation de l'espace sont dérivées à partir des deux cartes d'occupation du sol de 1987 et de 1999 qui ont été produites au chapitre précédent. Ces cartes comportent chacune 8 classes. L'opération de classification a été utilisée pour créer deux cartes d'utilisation de l'espace comportant 7 classes. Les classes d'occupation « savanes boisées » et « zones parcourues par le feu » correspondent à une seule utilisation de l'espace qui est appelée « brousse ». Les zones de montagne étant supposées stables au cours du temps, seules 6 classes sont utilisées dans les analyses. Une variable est dérivée pour chaque classe d'utilisation de l'espace et pour chaque date. Une variable correspondant aux changements d'utilisation de l'espace entre 1987 et 1999 est également calculée en croisant les deux premières cartes obtenues. Celle-ci indique pour chaque position si l'utilisation de l'espace présente a changé ou pas. On obtient un total de 18 variables binaires qui décrivent l'utilisation de l'espace et les changements survenus entre les deux dates (tableau 9.1).

Facteurs biophysiques

Les variables représentant les facteurs biophysiques concernent le relief, les types de sol et la pluviométrie (tableau 9.2). Les variables de relief (*altitude* et *slope*) représentent respectivement l'altitude et la pente. Le modèle numérique de terrain (MNT) est défini par la fonction f qui à chaque cellule de coordonnées x et y associe son altitude $f(x, y)$. La valeur de la variable *slope* est dérivée par la formule suivante où on identifie la valeur maximale de l'angle de changement d'altitude de chaque cellule aux cellules voisines :

$$P(x, y) = \text{Arctg} \left(\frac{\text{Max} \left(\frac{df(x, y)}{dx}, \frac{df(x, y)}{dy} \right)}{r} \right),$$

Cette valeur représente l'angle de déviation de la pente sur l'espace considérée. Par exemple, une pente de 45° correspondra à une valeur de 100%, ce qui signifie que la valeur de la pente pourrait excéder 100%.

Cette variable ne mesure pas la valeur réelle de la pente, elle n'est qu'un proxy de la pente qui dépend de la résolution du Modèle Numérique de terrain (MNT). Cette mesure aura tendance à sous estimer la pente en fonction de la variation de l'altitude. Pour améliorer le calcul de la pente à la résolution de 250 m, le MNT a été converti en une carte de point et une interpolation linéaire a été effectuée. La figure 9.4 présente la carte de la variable *altitude* où on peut voir que la zone d'étude est divisée en trois grandes zones d'altitude. D'est en ouest, on distingue successivement : les plaines du Diamaré où l'altitude varie entre 300 et 400 m, les zones de pédiplaines de Kaélé où l'altitude varie entre 400 m et 500 m et les zones

de piémonts où l'altitude varie entre 500 et 800 m. On note également quelques zones montagneuses, principalement dans la partie nord ouest qui atteignent environ 950 m.

Variable	Description	N_1	$\%N_1$
Utilisation de l'espace en 1987			
Habitation	1 si habitation, sinon 0	1 010	1,8
Maraîcher Verger	1 si maraîcher/verger, sinon 0	687	1,2
Culture pluviale	1 si culture pluviale, sinon 0	14 941	26,1
Sol nu	1 si sol nu, sinon 0	8 308	14,5
Sorgho repiqué	1 si sorgho de contre saison, sinon 0	10 276	18,0
Brousse	1 si brousse, sinon 0	19 045	33,3
Utilisation de l'espace en 1999			
Habitation	1 si habitation, sinon 0	1 068	1,9
Maraîcher Verger	1 si maraîcher/verger, sinon 0	794	1,4
Culture pluviale	1 si culture pluviale, sinon 0	20 490	35,9
Sol nu	1 si sol nu, sinon 0	5 007	8,8
Sorgho repiqué	1 si sorgho de contre saison, sinon 0	14 972	26,2
Brousse	1 si brousse, sinon 0	11 936	20,9
Changements d'utilisation de l'espace entre 1987 et 1999			
Habitation	1 si habitation a changé, sinon 0	62	0,1
Maraîcher Verger	1 si maraîcher/verger a changé, sinon 0	224	0,4
Culture pluviale	1 si culture pluviale a changé, sinon 0	7 641	13,4
Sol nu	1 si sol nu a changé, sinon 0	1 155	2,0
Sorgho repiqué	1 si sorgho de contre saison a changé, sinon 0	7 105	12,4
Brousse	1 si brousse a changé, sinon 0	2 219	3,9

Tableau 9.1 : Caractéristiques des variables décrivant l'utilisation de l'espace et ses changements. N_1 est le nombre d'observations où l'utilisation de l'espace concernée est présente et $\%N_1$ représente le pourcentage par rapport au nombre total d'observations $N=5714$.

Les variables de type de sol (*soilpot* et *soilsorg*) sont de type catégorique. Les catégories de ces variables correspondent à des niveaux d'aptitude croissants du sol pour les activités agricoles. Les valeurs de la variable *soilpot* correspondent à une appréciation des potentialités du sol pour l'ensemble des activités agricoles. Les caractéristiques de chaque niveau de potentialité sont données au chapitre 5. Les valeurs de la variable *soilsorg* correspondent à des niveaux d'aptitude des sols pour la culture du sorgho de contre saison. Les caractéristiques de chaque niveau d'aptitude ont été décrites au chapitre 7.

La variable de pluviométrie (*rainfall*) est calculée à partir des relevés pluviométriques disponibles dans la base de données sous forme de tableaux indiquant les quantités de pluies en mm et en nombre de jours de pluie pour 14 stations sur la zone d'étude pour la période 1994-2001. Une carte de points représentant les stations pluviométriques a été établie et une interpolation spatiale est effectuée pour dériver la pluviométrie annuelle moyenne de chaque cellule. La variable *rainfall* représentée spatialement à la figure 9.3, permet de distinguer quatre principales zones d'isohètes dont les valeurs de pluviométrie moyenne varient de 750 mm à 970 mm par an suivant un gradient nord est – sud ouest.

Variable	Description	Min.	Max.	Moy.	Ecart-type
Altitude	Altitude en mètre	331	968	441	82
Slope	Pente en % (45°=100%)	0	18	2	2
Rainfall	Pluviométrie annuelle moyenne en mm	751	973	864	46
Soilsorg	De 1 si très apte au sorgho repiqué à 6 si inapte	1	6	3	1,87
Soilpot	De 1 si très fortes potentialités agricoles à 6 si inapte	1	6	4	1,45

Tableau 9.2 : Caractéristiques des variables décrivant les facteurs biophysiques.

Facteurs sociodémographiques

Le tableau 9.3 donne la liste des variables représentant les facteurs sociodémographiques. Ces variables décrivent la pression potentielle de la population sur l'espace (*rurpop*, *poppot*), la densité de la population (*density*) et sa composition ethnique (*fulbé*, *guiziga* et *mofu*).

Variable	Description	Min.	Max.	Moy.	Ecart-type
Fulbe	Pourcentage de l'ethnie Foulbé	0	1	0,26	0,35
Mofu	Pourcentage de l'ethnie Mofu	0	1	0,23	0,38
Guiziga	Pourcentage de l'ethnie Guiziga	0	1	0,30	0,40
Density	Densité de la population	0	83 001	92	367
Rurpot87	Indice de pression de la population rurale en 1987	0	9 568	843	1 113
Rurpot99	Indice de pression de la population rurale en 1999	1	11 078	1 491	1 607
Poppot87	Indice de pression de la population totale en 1987	0	42 295	1 440	3 405
Poppot99	Indice de pression de la population totale en 1999	1	51 634	2 436	4 782

Tableau 9.3 : Caractéristiques des variables décrivant les facteurs sociodémographiques.

Les variables d'ethnicité, *fulbé*, *guiziga* et *mofu*, correspondent au pourcentage de chacun de ces groupes ethniques par localité. Une carte des villages associés aux effectifs de la population en 1987 a été réalisée sur la base des données du dernier recensement général de la population du Cameroun (RGPH, 1987). La population étant dynamique, une carte de population pour l'année 1999 a été réalisée en appliquant la formule de projection suivante : $P_n = P_{n_0} * (1 + T)^{n-n_0}$ où T est le taux d'accroissement de la population par unité administrative, n_0 est la date initiale et n est la date de projection.

La densité de population (*density*) a été calculée à partir de la carte des populations de villages en utilisant la méthode du polygone de Thiessen. Cette méthode d'interpolation locale suppose que la distribution de la population est homogène dans l'espace. Ce qui n'est pas très réaliste et ne pourrait pas permettre d'apprécier les différences d'utilisation de l'espace qui sont liées aux variations de la pression de la population. Pour résoudre ce problème un indice a été calculé pour représenter la pression potentielle de la population sur l'espace.

La formulation de cet indice s'appuie sur l'hypothèse selon laquelle la pression humaine sur l'espace et les ressources naturelles part d'un point central (le village) et évolue suivant un gradient décroissant quand on s'éloigne du village. Le modèle de calcul repose sur le principe général des modèles gravitaires qui définissent la relation physique entre objets, induite par une force de gravité en se basant sur la distance et une masse. Etant donné un objet (village) placé en un point origine (centre du village), la force (potentiel, pression) va diminuer graduellement au fur et à mesure qu'on se déplace de cette position centrale mais va également augmenter avec la masse (population) de l'objet. La fonction utilisée pour traduire ce facteur suit le modèle gaussien (Ingram, 1971) qui a été appliqué avec succès dans plusieurs études de cas traitant de l'accessibilité (Verburg et al, 2004). Cette fonction définit de la manière suivante produit un indice qui évolue suivant un gradient décroissant quand on s'éloigne des zones fortement peuplées.

$$Poppot_i = \sum_{j=1}^k \left[P_j * e^{\left(\frac{-(d_{ij})^b}{2a^2} \right)} \right]$$

Dans cette formule, $Poppot_i$ représente la pression sur la localité i , causée par les k villages les plus proches, d_{ij} est la distance entre cette localité i et le village j , a est la distance au point d'inflexion de la courbe Gaussienne que définit la fonction d'accessibilité et b est l'exposant de la distance. En s'appuyant sur les caractéristiques des villages de la zone d'étude (Balaza, Mowo, Gadas, voir chapitre 8), on a estimé la zone d'influence moyenne qui correspond à la distance au-delà de laquelle la pression de la population s'annule pour un village. Ainsi on a choisi $b = 2$ et le facteur distance a qui représente le point d'inflexion de la courbe Gaussienne à 1,5 km. La figure 9.3 montre que pour un village théorique de 1 500 habitants, l'indice de population décroît à partir de la valeur 1 500 et s'annule à 5 km du centre. Cet indice permet de localiser les zones de forte pression humaine et les zones relativement vides qui sont des meilleurs indicateurs de localisation des zones cultivées en cultures pluviales, en sorgho repiqué ou des zones encore boisées. Le modèle ainsi défini s'appliquerait donc bien à l'analyse de la structure spatiale de certaines utilisations de l'espace (cultures pluviales, sorgho repiqué et la brousse) dans les zones rurales où l'habitat dans les villages est regroupé. Ainsi, deux cas ont été distingués: 1) l'indice de population potentielle qui considère uniquement la population des zones rurales (*rurpot*); et 2) l'indice de population potentielle globale intégrant également la population urbaine (*poppot*). L'indice de pression de la population sur l'espace environnant a été calculé pour les deux dates afin de prendre en compte la dynamique de la population. A titre d'exemple, la carte de l'indice de population (*rurpot*) présentée à la figure 9.4, permet de voir les zones où la pression humaine dans les zones rurales est potentiellement plus forte. Ces zones qui apparaissent en couleur foncée sont principalement localisées dans la moitié nord de la région. Les zones plus claires dans la partie sud correspondent à des terres vides qui sont soit occupées par la brousse ou utilisées pour la culture du sorgho de contre saison.

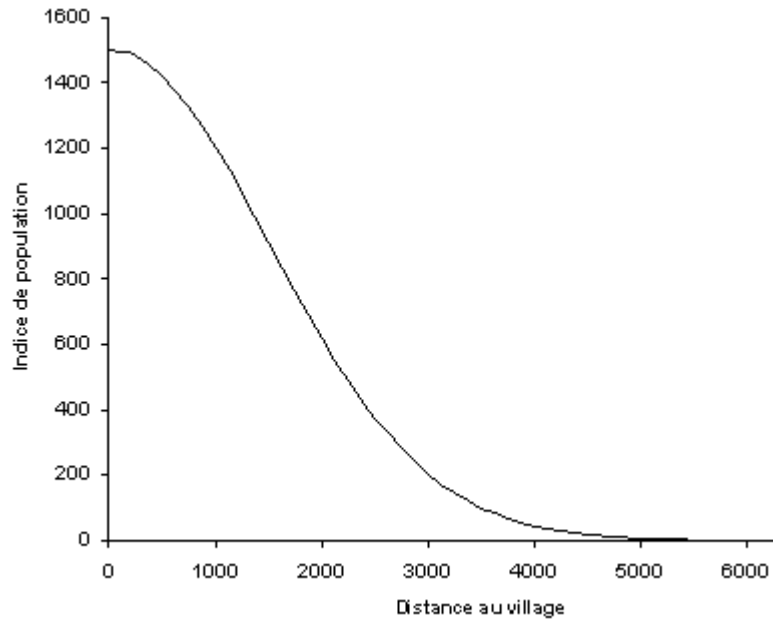


Figure 9.3 : Courbe de distribution de l'indice de population potentielle pour un village de 1 500 habitants.

Facteurs géoéconomiques

Les variables géoéconomiques décrites dans le tableau 9.4 concernent principalement les variables d'accessibilité. Il s'agit de la distance aux routes (*distroad*), la distance au cours d'eau (*distriver*), les temps estimés de déplacement à la ville (*travelc*), aux centres urbains (*travelu*) ou aux marchés (*travelm*) les plus proches. Ces variables ont été calculées en combinant plusieurs cartes de nature différentes qui sont disponibles dans la base de données de SMALL Savannah : une carte du réseau de l'infrastructure routière, une carte du réseau hydrographique, une carte de la ville de Maroua, une carte des petits centres urbains autour de Maroua et la carte des principaux marchés de sorgho et de bétail. La variable *livestock* dont les valeurs indiquent la densité du cheptel par zone est également considérée comme une variable économique dans la mesure où elle traduit l'importance de l'activité d'élevage. La variable *livestock* est calculée à partir des données de recensement du cheptel qui sont également disponibles dans la base de données de SMALL savannah. Les variables *distroad* et *distriver* représentant respectivement la distance à la route et la distance au cours d'eau le plus proche, sont obtenues en appliquant un calcul de distance euclidienne de chaque localité aux objets destination. Les cartes correspondant aux variables *distroad* et *distriver* présentées à la figure 9.4, sont faciles à comprendre. Elles mettent en exergue les réseaux routiers et hydrographiques de la zone d'étude ainsi que les gradients d'accessibilité évalués en terme de distance.

Les temps de déplacement relatif de chaque localité à la ville (*travelc*), au centre urbain (*travelu*) ou au marché (*travelm*) le plus proche ont été calculés en prenant en compte le niveau d'accessibilité. L'accessibilité est une fonction qui définit la capacité des populations à interagir avec des sites présentant des opportunités économiques comme le marché, les produits manufacturés, les moyens de transport, l'industrie etc. (Deichman, 1997). Plusieurs mesures de l'accessibilité peuvent être envisagées dépendant des acteurs, de leurs préférences et des conditions économiques et sociales (Verburg et al., 2004). En plus de la distance euclidienne entre la localité et le point destination, la fonction d'accessibilité définie dans cette application prend en compte les contraintes de déplacement posées par les moyens de déplacements

utilisés dans la région, le retard crée par la présence d'un cours d'eau ou la nature des différents types d'utilisation de l'espace. Ces contraintes sont modélisées par un facteur de résistance (Map_{Resist}) qui est inversement proportionnelle à la vitesse et traduit un poids ou qui reflète la diminution relative de la vitesse dans les zones moins accessibles comparée aux zones les mieux accessibles (Toxoepus, 1996).

$$Map_{Resist} = \frac{Map_1}{Map_{SpeedRoad} \oplus Map_{SpeedRiver} \oplus Map_{SpeedLU}} \text{ où,}$$

Map_1 est une carte avec toutes les valeurs à 1, $Map_{SpeedRoad}$, $Map_{SpeedRiver}$ et $Map_{SpeedLU}$ représentent respectivement les vitesses de déplacement affectées aux routes, aux rivières et aux différentes unités d'occupation du sol. Ces vitesses ont été estimées à partir des connaissances de la région issues de la littérature et des travaux de terrain.

Variable	Description	Min.	Max.	Moy.	Ecart-type
livestoc	Densité du cheptel en UBT/km ²	0	98	52	17
travelc	Temps de déplacement à la ville de Maroua	0	33 639	14034	5 964
travelu	Temps de déplacement au centre urbain plus proche	0	25 063	9053	4 688
travelm	Temps de déplacement au marché plus proche	0	32 205	10569	5 792
distriver	Distance à la rivière en mètre	0	14 160	3875	3 051
distroad	Distance à la route en mètre	0	7603	1559	1 290

Tableau 9.4 : Caractéristiques des variables décrivant les facteurs géoéconomiques.

Pour les routes, les vitesses moyennes estimées prennent en compte le type de route et le moyen de transport utilisé. Nous avons distingué cinq moyens différents de déplacement et la vitesse moyenne pour chaque catégorie de route est une moyenne pondérée qui prend en compte les vitesses relatives de chaque moyen de déplacement et le pourcentage des utilisateurs qui l'empruntent (tableau 9.5).

Pour les cours d'eau, l'estimation s'est basée sur une combinaison des données de la littérature et des connaissances de la zone d'étude. Dans plusieurs applications où les cours d'eau contiennent de l'eau, la vitesse de l'eau est placée à une valeur faible, de l'ordre de 0,3 km/h, traduisant ainsi les difficultés d'accessibilité (Toxoepus, 1996 ; Verburg, 2004). Dans le cas du Nord Cameroun, la contrainte des cours d'eau est très limitée dans l'espace et le temps car ce sont des cours d'eau saisonniers qui ne contiennent de l'eau que trois mois dans l'année et peuvent facilement être traversés à pied. La vitesse de 0,5 km/h a été affectée aux cours d'eau principaux et la vitesse de 1 km/h aux cours d'eau secondaires.

	Type de route	Moyen de déplacement					Vitesse Moyenne	Facteur de résistance
		Voiture Privée	Voiture publique	Moto	Bicyclette	Pied		
	Bitumée	80	60	25	15	5	51	1
	Non bitumée et permanente	40	30	15	10	4	26	1,96
	Non bitumée et non permanente	27	20	10	7	3	18	2,83
	Piste	15	10	8	3	2	9	5,60
	Pourcentage des utilisateurs	5	70	15	5	5		

Tableau 9.5 : Vitesse de déplacement moyenne en km/h pour les différentes catégories de route. Les valeurs de vitesse sont des estimations basées sur les expériences de terrain et les avis des utilisateurs et experts de la région.

La carte d'occupation du sol sur la zone d'étude a été utilisée pour regrouper les différentes unités d'occupation du sol en fonction des difficultés d'accès et leur affecter des vitesses correspondantes. Ainsi, les zones montagneuses sont plus difficilement accessibles (0,25 km/h). Les zones de brousses sont également relativement difficiles à accéder du fait de la présence de la végétation ligneuse (2 km/h). Par contre les sols nus et les zones cultivées sont plus facilement accessibles. Les zones de culture de sorgho de contre saison sur les terres argileuses présentent toutefois des difficultés d'accès en saison des pluies (3 km/h) alors que les zones sur les terres sableuses ou caillouteuses sont relativement accessibles (5 km/h).

Le facteur de résistance étant inversement proportionnel à la vitesse et le facteur de résistance associé à la plus grande vitesse étant fixée à 1, le facteur de résistance associé aux autres vitesses de déplacement est déduit facilement comme le présente le tableau 9.6. Le temps de déplacement ($Travel_{dest}$) est évalué comme une distance pondérée en utilisant une fonction d'analyse spatiale qui prend en entrée la carte de résistance et une carte de destination (Map_{Dest}). Cette dernière peut être la carte de la ville, des centres urbains ou des marchés. On note :

$Travel_{Dest} = Weightdist(Map_{Dest}, Map_{Resist})$. Dans ce calcul, plus le poids d'une cellule c'est-à-dire la valeur du facteur de résistance sera élevée et plus la distance sera grande, c'est-à-dire que le temps pour accéder à la destination sera plus important.

A titre d'exemple, la représentation spatiale de la variable *travelu* est donnée à la figure 9.4. Les zones les plus claires sont les plus accessibles à partir des centres urbains. On peut y distinguer 4 principaux centres urbains : le principal centre urbain situé au cœur de la région correspond à la ville de Maroua est entouré des centres de Méri dans le nord ouest, de Gazawa dans le sud ouest et de Mindif dans le sud est. L'influence du centre urbain de Bogo sur l'accessibilité est également perceptible dans la partie nord est de la région. Les zones les moins accessibles apparaissent en couleur marron foncé.

N°	Type d'occupation du sol	Vitesse moyenne en km/h	Facteur de résistance
1	Route Bitumée	51	1
2	Route non bitumée permanente	26	1,96
3	Route non bitumée et Non permanente	18	2,83
4	Piste	9	5,6
5	Zones sableuses	5	16
6	Zones argileuses	3	26
7	Brousse	2	40
8	Cours d'eau secondaire	1	80
9	Cours d'eau principal	0,5	160
10	Montagnes	0,25	320

Tableau 9.6 : Facteur de résistance associé aux différentes entités spatiales.

9.4. Résultats de l'exploration des facteurs déterminants et construction des modèles de prédiction

Cette section présente et interprète tour à tour les résultats de la procédure d'exploration des facteurs déterminant chacune des 6 utilisations de l'espace suivantes : zone d'habitation, culture maraîchère/verger, culture pluviale, sol nu, sorgho repiqué et brousse.

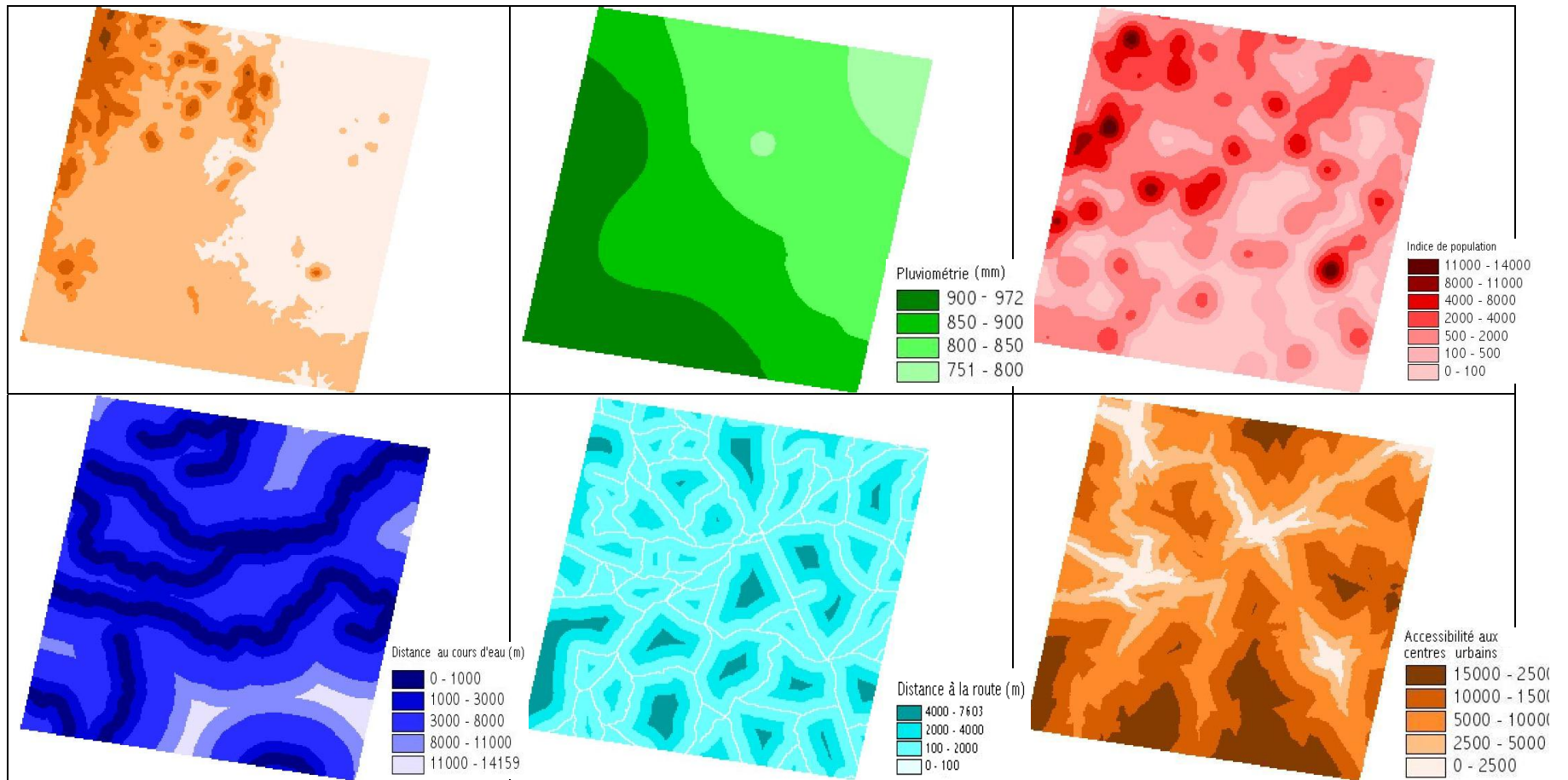


Figure 9.4 : Cartes de quelques variables représentant les facteurs déterminant la structure et les changements d'utilisation de l'espace.

9.4.1. Les zones d'habitation

Le tableau 9.7 résume les résultats des régressions simples entre l'utilisation de l'espace « zone d'habitation » et l'ensemble des facteurs potentiellement déterminants dans les trois cas étudiés soit 1987, 1999 et le changement entre les deux dates.

Variable	1987		1999		Changement	
	β	ROC	β	ROC	β	ROC
Altitude	-0,003	0,553*	-3×10^{-3}	0,561*	-0,003	0,538***
Slope	+0,072	0,535*	-0,156	0,535**	+0,172	0,634
Soilpot		0,655		0,663		0,814
Soilpot1	+0,836		+0,905		+1,447	
Soilpot2	-0,744		-0,744		-1,794	
Soilpot3	-0,250		-0,250		-1,418	
Soilpot5	-0,317		-0,271		-1,525	
Soilpot6	+0,182		+0,362		+1,170	
Soilsorg		0,653		0,661		0,790
Soilsorg1	-1,189		-1,227		-0,059	
Soilsorg2	-0,326		-0,372		-13,618	
Soilsorg3	-0,094		-0,140		-0,452	
Soilsorg5	+1,065		+1,087		+2,072	
Soilsorg6	+0,157		+0,216		+0,991	
Rainfall	-0,014	0,684	-0,014	0,690	-0,021	0,787
Densite	+0,0014	0,771	+0,001	0,763	$+4,9 \times 10^{-4}$	0,587
Rurpot	$+3,8 \times 10^{-4}$	0,705	$+1,4 \times 10^{-4}$	0,705	$+2 \times 10^{-5}$	0,651***
Poppot	$+1,9 \times 10^{-4}$	0,823	$+2,7 \times 10^{-4}$	0,815	$+0,7 \times 10^{-5}$	0,845
Fulbé	+1,06	0,595	+1,142	0,603	+2,295	0,738
Guiziga	-1,462	0,600	-1,502	0,603	-2,014	0,659
Mofu	-0,782	0,553*	-0,793	0,552	-1,597	0,580*
Distriver	$-1,6 \times 10^{-4}$	0,699	$-1,8 \times 10^{-4}$	0,680	-0,001	0,810
Distroad	-0,01	0,775	-0,001	0,776	-0,001	0,676
Travelc	$-2,2 \times 10^{-4}$	0,762	$-2,2 \times 10^{-4}$	0,770	$-3,1 \times 10^{-4}$	0,870
Travelu	$-2,4 \times 10^{-4}$	0,737	$-2,4 \times 10^{-4}$	0,739	$-2,2 \times 10^{-4}$	0,740
Travelm	$-1,7 \times 10^{-4}$	0,710	$-1,8 \times 10^{-4}$	0,716	$-2,6 \times 10^{-4}$	0,783
Livestock	-0,083	0,701	-0,086	0,709	-0,115	0,873

Tableau 9.7 : Paramètres des régressions simples pour l'utilisation de l'espace « habitation ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, * pour $0,1 < P < 0,9$.**

On peut remarquer dans le tableau 9.7 que les trois variables *altitude*, *slope* et *mofu* qui représentent respectivement le relief, et l'ethnicité affichent des corrélations non significatives ($P > 0.01$). Les variables d'ethnicité *fulbe* et *guiziga* quant à elles ne fournissent pas une explication suffisante ($ROC < 0,6$). Les variables *travelc*, *travelu*, *travelm* et *livestock* relatives aux centres urbains et aux marchés qui affichent par contre un pouvoir explicatif supérieur à 0,7 n'ont pas été sélectionnées parce qu'elles ne traduisent pas un processus connu en relation avec l'installation des zones d'habitation. La relation avec les centres urbains traduit tout simplement la concentration des habitations au niveau des pôles d'urbanisation. C'est également le cas pour les variables d'ethnicité (*fulbé*) et de pluviométrie (*rainfall*), qui sont fortement corrélées avec les changements mais pour lesquelles il est difficile d'associer un processus pertinent.

Les variables de population (*density*, *poppot*, *rurpot*) affichent les pouvoirs explicatifs les plus élevés ($ROC = 0,77$), ce qui exprime bien l'évidence empirique selon laquelle la population explique mieux l'importance et la distribution des zones d'habitation. De la même façon, les pouvoirs explicatifs des variables *distriver* et *distroad* traduiraient le fait que l'installation des zones d'habitation s'effectue prioritairement le long des routes et des cours d'eau. Les corrélations avec les variables de type de sol (*soilsorg* et *soilpot*) sont significatives, mais affichent des pouvoirs explicatifs pas assez élevés (resp. 0,655 et 0,653). Toutefois, les observations de terrain montrent que les terres argileuses qui sont des sols très aptes à la culture du sorgho de contre saison ne sont pas utilisées pour construire les habitations. La variable *soilsorg* a été sélectionnée et pourrait donc mieux expliquer la distribution spatiale des zones d'habitation que la variable *soilpot* qui affiche pourtant un pouvoir explicatif plus élevé.

Pour la situation de 1987, les 6 variables *soilsorg*, *distroad*, *distriver*, *density*, *poppot* et *rurpot* ont été insérées dans la régression conditionnelle. Les variables *density*, *distriver* et *rurpot* ont été éliminées à l'issue de cette procédure. Le modèle de régression final obtenu comprend les trois variables *poppot*, *distroad* et *soilsorg* qui expliquent mieux la distribution des zones d'habitation. On peut voir dans le tableau 9.8 qui résume les coefficients logistiques que la probabilité d'installation des zones d'habitation sera plus grande dans les zones où, la population est importante et de préférence à proximité des axes routiers. Le signe de coefficients des variables *soilsorg1* et *soilsorg2* confirme l'hypothèse et les observations faites sur les difficultés à construire sur des sols argileux. La capacité de discrimination du modèle de prédiction final obtenu est très bonne ($ROC = 0,886$). Comme le montre la figure 9.5, la superficie sous la courbe ROC est suffisamment importante. Cette superficie variant entre 0 et 1. Les valeurs proches de 1 comme dans ce cas indiquent que le test de discrimination est presque toujours vrai. Un modèle aléatoire donnerait une valeur proche de 0,5 et la courbe se rapprocherait de la diagonale. Pour la situation de 1999, la même combinaison de variables a été insérée dans la régression conditionnelle. On obtient les mêmes variables explicatives mais la capacité de discrimination du modèle obtenu est légèrement plus élevée ($ROC = 0,887$). Le signe de tous les coefficients coïncident entre les deux dates et seules leurs valeurs sont différentes.

Pour le cas des changements entre 1987 et 1999, la même combinaison de variables a été insérée dans la régression conditionnelle. Les mêmes variables sont rejetées à l'exception de *distriver* et le modèle de prédiction des changements obtenu affiche une plus grande capacité de détection des changements ($ROC = 0,889$). Les coefficients des variables *poppot* et *soilsorg* qui sont conservées ne traduisent pas un processus pertinent par rapport au changement des zones d'habitation. Ils traduisent en réalité le fait que

les zones d'habitations qui ont changé appartiennent à des zones densément peuplées et de préférence sur les sols ayant un faible taux d'argile comme le montre les deux modèles précédents.

Toutefois, l'apparition de la variable *distriver* et la faiblesse relative du niveau de signification de la variable *distroad* ($0,01 < P < 0,05$) traduisent le fait que la probabilité de déplacement des zones d'habitation augmente avec la proximité des cours d'eau dont les débordements en saison de pluies exposent les populations à des risques. Cette dynamique a été observée le long du cours d'eau Mayo Tsanaga, à partir des images satellites et confirmée sur le terrain à l'échelle des terroirs. Dans le terroir de Mowo, par exemple, les populations qui se sont déplacées se sont réinstallées le long de la route. En effet les changements de l'utilisation « habitation » concernent de très faibles superficies qui rendent l'analyse difficile à l'échelle régionale. On observe des dynamiques plus marquées à l'échelle du terroir où des facteurs déterminants plus pertinents se révéleraient à cette échelle.

1987		1999		Changement	
Variables	β_i	Variables	β_i	Variables	β_i
β_0	- 3,603	β_0	- 3,57998	β_0	- 5,88411
Poppot87	+ 0,00018	Poppot99	+ 0,00013	Poppot99	+ 0,00004
Distroad	- 0,00119	Distroad	- 0,00119	Distroad	- 0,00040*
Soilsorg1	- 0,04794	Soilsorg1	- 0,78685	Distriver	- 0,00043
Soilsorg2	- 0,06710	Soilsorg2	- 0,02660	Soilsorg1	- 0,30626
Soilsorg3	+ 0,96528	Soilsorg3	+ 0,28827	Soilsorg2	- 13,4090
Soilsorg5	+ 0,54001	Soilsorg5	+ 0,40248	Soilsorg3	- 0,67566
Soilsorg6	- 0,01984	Soilsorg6	- 0,17577	Soilsorg5	+ 0,96856
				Soilsorg6	+ 0,50159
ROC = 0,886		ROC = 0,887		ROC = 0,889	

Tableau 9.8 : Paramètres des modèles de régression finaux de l'utilisation de l'espace « zone d'habitation ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, *** pour $0,1 < P < 0,9$.

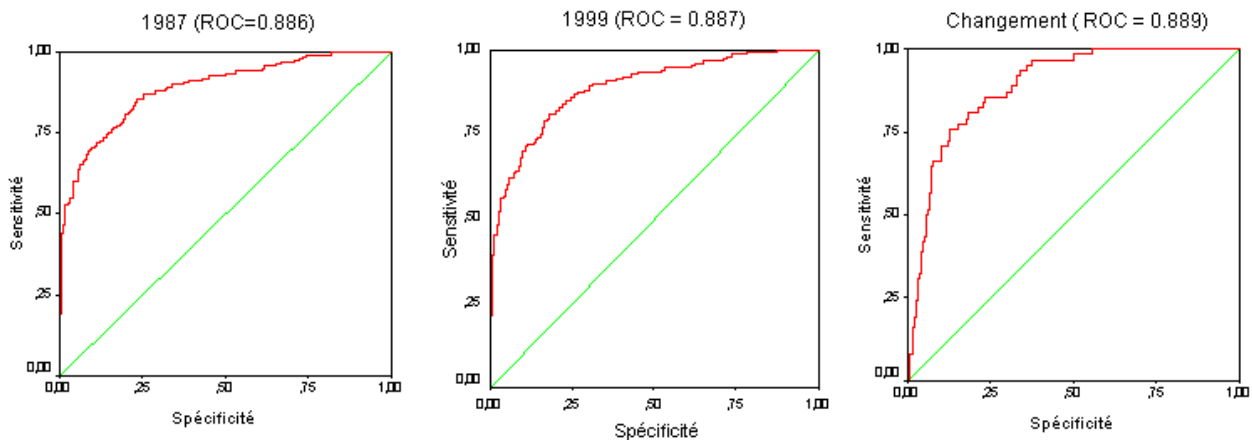


Figure 9.5 : Courbes ROC des modèles de régression finaux de l'utilisation de l'espace « habitation ».

9.4.2. Les zones de vergers ou cultures maraîchères

Le tableau 9.9 présente les résultats des régressions simples, qui sont utilisés pour la construction du modèle de prédiction de la distribution spatiale de l'utilisation de l'espace « maraîcher/verger ».

Variable	1987		1999		Changement	
	β	ROC	β	ROC	β	ROC
Altitude	-0,012	0,665	-0,008	0,604	-0,005	0,556
Slope	-0,117	0,561*	-0,156	0,550*	-0,465	0,588
Soilpot		0,748		0,749		0,724
Soilpot1	+3,623		+3,703		+16,827	
Soilpot2	+1,563		+1,927		+15,360	
Soilpot3	+0,627		+0,627		+14,195	
Soilpot4	+1,617		+1,809		+15,576	
Soilpot5	+1,745		+1,049		+13,946	
Soilsorg		0,770		0,766		0,702
Soilsorg1	-1,852		-1,005		-0,092	
Soilsorg2	+0,493		+0,554		+0,002	
Soilsorg3	+0,629		-0,033		-0,275	
Soilsorg4	+2,226		+2,053		+1,457	
Soilsorg5	-0,251		-0,727		-1,048	
Rainfall	-0,012	0,665	-0,007	0,598	-0,001	0,517***
Density	+4,6x10 ⁻⁴	0,694	+4,7x10 ⁻⁴	0,725	+2,6x10 ⁻⁴	0,687*
Rurpot	+3,4x10 ⁻⁴	0,716	+0,6x10 ⁻⁴	0,735	+2x10 ⁻⁴	0,691
Poppot	+0,9x10 ⁻⁴	0,751	+2,3x10 ⁻⁴	0,780	+0,7x10 ⁻⁴	0,752
Fulbé	+1,658	0,691	+1,386	0,656	+0,220	0,528***
Guiziga	-1,507	0,570	-0,764	0,521*	+0,195	0,559***
Mofu	-0,760	0,507*	-0,361	0,482***	-0,074	0,466***
Distriver	-15,3x10 ⁻⁴	0,912	-0,001	0,892	-9,35x10 ⁻⁴	0,861
Distroad	-0,001	0,690	-0,001	0,665	-2,9x10 ⁻⁴	0,580
Travelc	-1,64x10 ⁻⁴	0,727	-1,5x10 ⁻⁴	0,723	-0,9x10 ⁻⁴	0,652
Travelu	-1,44x10 ⁻⁴	0,667	-1,4x10 ⁻⁴	0,667	-0,8x10 ⁻⁴	0,590
Travelm	-1,2x10 ⁻⁴	0,667	-1,5x10 ⁻⁴	0,700	-1,3x10 ⁻⁴	0,669
Livestock	-3,1x10 ⁻⁴	0,486***	+0,006	0,542***	+0,012	0,552*

Tableau 9.9 : Paramètres des régressions simples pour l'utilisation de l'espace « Maraîcher/verger ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, *** pour $0,1 < P < 0,9$.

Le tableau 9.9 affiche quatre variables avec des corrélations non significatives. Il s'agit des variables *livestock*, *slope*, *guiziga* et *mofu*. Les variables de relief (*altitude*), de pluviométrie (*rainfall*), d'ethnicité (*fulbé*) et la variable *distroad* ont un pouvoir explicatif relativement élevé ($0,65 < \text{ROC} < 0,7$), mais ne sont pas sélectionnées du fait qu'elles ne traduisent ni un processus connu, ni une hypothèse pertinente à tester. La corrélation assez élevée avec la variable d'ethnicité (*fulbé*) s'explique par le fait que l'aire de distribution des populations Foulbé (plaine du Diamaré) coïncide avec les zones de cultures maraîchères/vergers.

Les variables suivantes affichent les meilleurs pouvoirs explicatifs ($\text{ROC} > 0,7$) : *distriver*, *soilpot*, *soilsorg*, *density*, *rurpot*, *poppot*, *travelc*. De plus chacune des corrélations avec ces variables traduit un processus pertinent ou une hypothèse à tester. Elles ont par conséquent été sélectionnées pour la régression conditionnelle. La sélection des variables *distriver*, *soilpot*, et *soilsorg* se justifie par les connaissances disponibles sur cette utilisation de l'espace dans la région. On a vu au chapitre 5 que la distribution spatiale des cultures maraîchères/vergers est d'abord déterminée par les conditions du milieu notamment le type de sol et la proximité des cours d'eau. Les observations sur les images satellites et les visites de terrain montrent que la culture se pratique principalement le long des cours d'eau ou sur les alluvions des lits majeurs des cours d'eau où la nappe phréatique est peu profonde et donc facilement accessible. C'est le cas par exemple dans la plaine du Diamaré où les maraîchers sont concentrés le long des cours d'eau mayo Tsanaga et mayo Boula et dans la plaine de Koza à Mora, le long des cours d'eau temporaires mayo Ngassawé et mayo Moskota. La sélection des variables de population *density*, *poppot* et *rurpot* vise à tester l'influence de la croissance démographique sur l'expansion de la culture. Les résultats pourraient fournir des bases pour apprécier et discuter la prévalence des théories de Boserup ou de Malthus (chapitre 2). Par contre, les variables *travelu*, et *travelm* affichent un pouvoir explicatif relativement faible (0,6 et 0,7). Toutefois, elles ont été sélectionnées pour tester l'hypothèse selon laquelle le développement récent de l'agriculture maraîchère est favorisé par la proximité des petits centres urbains qui constituent un marché potentiel et une source d'investissement dans la production.

Pour la situation de 1987, les 9 variables suivantes ont été insérées dans la procédure de régression conditionnelle pas à pas : *distriver*, *travelc*, *travelu*, *travelm*, *density*, *poppot*, *rurpot*, *soilsorg* et *soilpot*. Les variables *soilpot* et *density* ont été exclues de l'équation du modèle final. Ceci nous amènerait à conclure que la densité de la population n'explique pas la distribution spatiale des maraîchers/vergers et que la variable *soilsorg* apparaît plus explicative que *soilpot*. On peut voir dans le tableau 9.10 que les coefficients des variables *distriver*, *travelc* et *travelu* confirment l'évidence selon laquelle les zones de cultures maraîchères se développent de préférence à proximité des cours d'eau et principalement autour des centres urbains qui assurent l'investissement nécessaire. Toutefois, il est a priori difficile d'expliquer raisonnablement la relation avec la variable *travelm*. L'explication c'est que les marchés considérés dans la définition de cette variable ne concernent pas principalement les produits maraîchers et fruitiers, mais plutôt les marchés de bétail et de sorgho de contre saison. Les produits maraîchers et fruitiers sont principalement orientés vers les marchés des centres urbains. La capacité de discrimination du modèle obtenu est la meilleure de toutes les utilisations de l'espace ($\text{ROC} = 0,939$). On peut voir sur la figure 9.6 que la courbe ROC se rapproche de la situation idéale où $\text{ROC}=1$.

Pour la situation de 1999, la même combinaison de variables que celle de 1987 a été insérée dans la procédure de régression conditionnelle. Malgré l'importance de la variable *density* dans la régression

simple (ROC = 0,725), celle-ci a été exclue de l'équation du modèle final. Les variables *soilpot* et *travelu* ont été également exclues. Les coefficients logistiques du modèle obtenu gardent les mêmes signes que pour la situation de 1987. La capacité de discrimination du modèle est légèrement plus faible mais reste très élevée soit ROC = 0,915.

Pour la carte des changements entre 1987 et 1999, la même combinaison de variables a été sélectionnée. En plus de la variable *density*, les variables *soilpot* et *poppot* ont été exclues de l'équation du modèle. La capacité de détection des changements de cette utilisation de l'espace par le modèle obtenu est évaluée à 0,888. On remarque à partir des coefficients du modèle obtenu que les variables *distriver*, *travelc* expliquerait la distribution des changements de l'utilisation de l'espace « maraîcher/verger ». Le signe du coefficient de la variable *travelu* est par contre difficile à associer avec une relation pertinente traduisant les changements observés. Comme dans le cas des zones d'habitation, ces relations traduisent plutôt la localisation de cette utilisation de l'espace qui a été mis en évidence dans les deux modèles statiques précédents. En effet, les changements observés pour l'utilisation de l'espace « maraîcher/verger » sont également de très faibles superficies et donc difficiles à expliquer à l'échelle régionale. On a vu au chapitre 8 que ces espaces sont convertis principalement pour les cultures pluviales ou le sorgho repiqué. Les observations de terrain ont montré que ce type de changement s'effectue en général sur les alluvions du lit majeur des cours d'eau où les sols peuvent se prêter à plusieurs types de culture. Le type de conversion d'une année à l'autre dépend de la pluviométrie et des stratégies paysannes pour le maintien de la sécurité alimentaire.

1987		1999		Changement	
Variables	β_i	Variables	β_i	Variables	β_i
Bo	-0,86642	Bo	-1,39035	Bo	-3,51716
Distriver	-0,00150	Distriver	-0,00108	Distriver	-0,00083
Travelc	-0,00033	Travelc	-0,00023	Travelc	-0,00009
Travelu	-0,00017	Travelm	0,00012	Travelu	+0,00027
Travelm	+0,00034	Rurpop99	0,00034	Travelm	-0,00019
Rurpop87	+0,00053	Poppot99	-0,00004	Rurpop99	+0,00022
Poppot87	-0,00006	Soilsorg1	-1,44444	Soilsorg1	-0,61235
Soilsorg1	-2,19931	Soilsorg2	+0,83701	Soilsorg2	-0,06472
Soilsorg2	0,94809	Soilsorg3	-0,80407	Soilsorg3	-0,54877
Soilsorg3	-0,69117	Soilsorg5	0,29364	Soilsorg5	-0,15109
Soilsorg5	+0,32527	Soilsorg6	-1,78467	Soilsorg6	-1,56808
Soilsorg6	-1,75173				
ROC = 0,939		ROC = 0,915		ROC = 0,888	

Tableau 9.10 : Paramètres des modèles de régression finaux de l'utilisation de l'espace « maraîcher/verger ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, *** pour $0,1 < P < 0,9$.

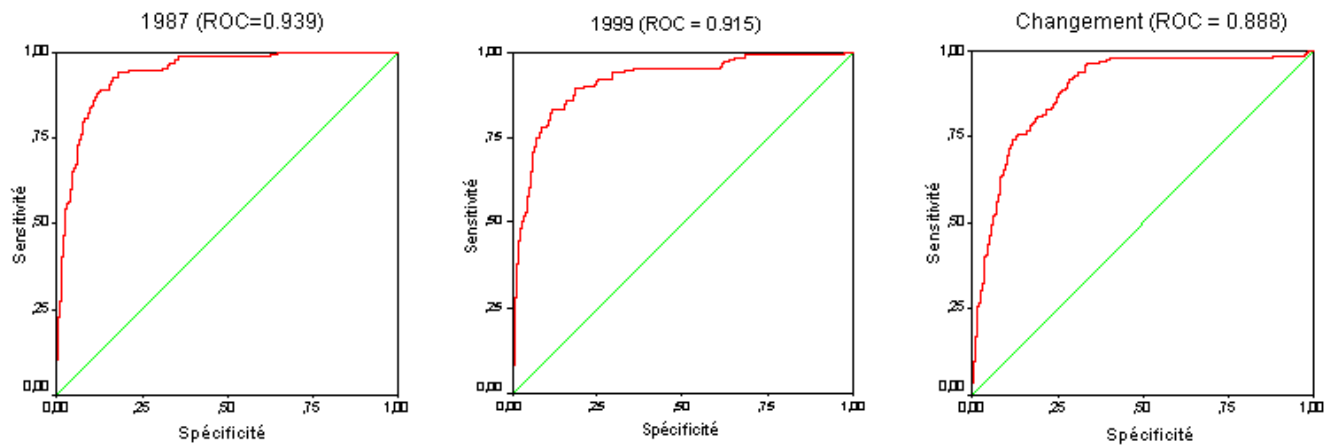


Figure 9.6 : Courbes ROC des modèles de régression finaux de l'utilisation de l'espace « Maraîcher/verger ».

9.4.3. Les zones de culture pluviale

Le tableau 9.11 résume les résultats des régressions simples qui sont utilisés pour la construction du modèle de distribution des zones de culture pluviale.

Variable	1987		1999		Changement	
	β	ROC	β	ROC	β	ROC
Altitude	+0,002	0,537	+0,004	0,596	+0,002	0,593
Slope	+0,106	0,565	+0,137	0,605	+0,02736	0,572
Soilpot		0,654		0,661		0,621
Soilpot1	+0,674		+0,447		-0,078	
Soilpot2	- 0,890		-1,152		-0,701	
Soilpot3	- 0,532		-0,178		+0,473	
Soilpot4	+ 0,130		+0,218		+0,427	
Soilpot5	+ 0,372		+0,371		-0,024	
Soilsorg		0,660		0,671		0,630
Soilsorg1	- 0,538		-0,360		-1,832	
Soilsorg2	- 0,729		-1,846		-0,997	
Soilsorg3	+ 0,175		-1,448		-0,683	
Soilsorg4	+ 0,920		-0,857		-0,582	
Soilsorg5	+ 0,612		+0,050		-0,369	
Rainfall	- 2,9x10 ⁻³	0,547	+2x10 ⁻³	0,524	+0,00721	0,601
Densite	+ 2,2x10 ⁻⁴	0,635	+7x10 ⁻⁵	0,609***	-2,6x10 ⁻⁴	0,483
Rurpot	+ 3,7x10 ⁻⁴	0,674	+1,5x10 ⁻⁵	0,644	+0,8x10 ⁻⁴	0,528
Poppot	+ 4,9x10 ⁻⁵	0,675	+2,19x10 ⁻⁴	0,632	-0,23x10 ⁻⁴	0,482
Fulbé	- 0,216	0,509	-0,807	0,564	-1,039	0,587
Guiziga	- 0,478	0,530	-0,299	0,526	+0,065	0,492*
Mofu	+ 0,943	0,605	+1,333	0,626	+0,774	0,561
Distriver	- 1,25x10 ⁻⁴	0,622	-5x10 ⁻⁵	0,564	+0,6x10 ⁻⁴	0,553
Distroad	- 3,9x10 ⁻⁴	0,627	-3x10 ⁻³	0,611	-0,4x10 ⁻⁴	0,521
Travelc	- 8x10 ⁻⁵	0,626	-5x10 ⁻⁵	0,569	+0,2x10 ⁻⁴	0,550
Travelu	-9,9x10 ⁻⁵	0,623	-9x10 ⁻⁵	0,616	-0,3x10 ⁻⁴	0,530
Travelm	-7x10 ⁻⁵	0,601	-5x10 ⁻⁵	0,572	-3,7x10 ⁻⁶	0,493**
Livestock	5,4x10 ⁻³	0,512	-0,001	0,522***	-0,005	0,539

Tableau 9.11 : Paramètres des régressions simples pour l'utilisation « Culture pluviale ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, * pour $0,1 < P < 0,9$.**

Dans l'ensemble, on trouve qu'aucun facteur n'explique clairement la distribution spatiale des zones de culture pluviale. On peut remarquer dans le tableau 9.11 que deux variables notamment *livestock* et *density* affichent des corrélations non significatives et ne peuvent par conséquent être sélectionnées pour la régression conditionnelle. Il en est de même pour la plupart des variables comme celles du relief (*altitude*, *slope*), de la pluviométrie (*rainfall*), et d'ethnicité (*fulbé*, *Guiziga*) qui affichent un pouvoir explicatif inférieur à 0,6. La variable d'ethnicité (*mofu*) et Les variables géoéconomiques (*distriver*, *distroad*, *travelc*, *travelu* et *travelm*) affichent une corrélation significative et un pouvoir explicatif assez grand ($0,6 < ROC < 0,7$) mais n'ont pas pu être associées à un processus pertinent ou connu. La corrélation avec la variable *mofu* correspond au fait que l'aire de distribution de l'ethnie Mofu est localisée en zone des piémonts qui est une zone où le système agraire est dominé par les cultures pluviales.

Par contre, les variables *soilsorg*, *soilpot*, *poppot* et *rurpop* qui affichent des corrélations significatives avec des pouvoirs explicatifs compris entre 0,6 et 0,7 ont été sélectionnées pour la procédure de régression conditionnelle. Comme pour les autres formes d'utilisation de l'espace, la densité (*density*) de population n'apparaît pas comme un facteur explicatif pertinent de la distribution spatiale des zones de cultures pluviales. L'indice de pression de la population qui est représentée par les variables *poppot* et *rurpot* explique mieux la distribution des zones cultivées en culture pluviale.

Pour la situation de 1987, seules les variables de type de sol et de population ont été insérées dans la procédure de régression conditionnelle au vu des observations précédentes. A l'issue de cette procédure, les variables *poppot* et *density* sont exclues de l'équation du modèle final qui affiche une capacité de discrimination égale à 0,693. La variable *soilsorg* affiche un niveau de signification plus grand que *soilpot* mais, en suivant une approche déductive, on privilégie la variable *soilpot* qui devrait mieux expliquer la distribution des zones de culture pluviale et on obtient un modèle de régression final avec une capacité de discrimination de 0,695. Le modèle où la variable *soilsorg* est incluse affiche une performance de 0,694. Les coefficients logistiques du modèle obtenu traduisent bien les hypothèses formulées. En effet, la probabilité pour que l'utilisation de l'espace soit de type « culture pluviale » est élevée dans les zones où la pression humaine est forte. Compte tenu de la représentation spatiale de l'indice de pression humaine et considérant l'hypothèse d'une organisation de l'utilisation de l'espace en auréoles concentriques autour des zones d'habitations (chapitre 7), les zones de culture pluviale constituent donc effectivement la première auréole autour des zones d'habitation où les valeurs de l'indice sont élevées. Cette tendance d'organisation de l'utilisation de l'espace a été observée par la cartographie de l'occupation du sol à l'échelle des terroirs et de la petite région autour de Maroua (chapitres 7 et 8).

Pour la situation de 1999, le même ensemble de variables a été sélectionné pour la régression conditionnelle et les résultats obtenus sont similaires. Le modèle de régression obtenu affiche une capacité de discrimination évaluée à 0,695. Comme le montre le tableau 9.12, le signes des coefficients logistiques coïncident avec ceux de la situation de 1987 et seule les valeurs sont différentes.

Pour la carte des changements entre 1987 et 1999, les variables de population affichent un pouvoir explicatif très bas. Seules les variables de sol (*soilsorg* et *soilpot*) et de pluviométrie (*rainfall*) ayant un pouvoir explicatif supérieur à 0,6 ont été sélectionnées pour la régression conditionnelle. Elles ont été toutes maintenues et le modèle obtenu à une capacité de détection des changements évaluée à 0,666. En maintenant uniquement les variables *soilpot* et *rainfall*, le modèle final obtenu a une capacité de

discrimination de 0,665. Les changements observés pour l'utilisation de l'espace culture pluviale sont les plus importants mais de nature assez diverse, ce qui rend difficile l'explication. Les principaux changements de cette classe s'effectuent pour les classes « sorgho de contre saison » et « sol nu ». Les autres changements concernent la conversion en « zones brousse » (jachères) et en cultures maraichères/vergers. Une très faible proportion est convertie en zone d'habitation. Le modèle obtenu prédit que la probabilité de changement de cette utilisation est d'autant plus grande que la pluviométrie est élevée. Nous n'avons pas pu trouver une explication raisonnable à cette corrélation d'autant plus que cette pluviométrie est une moyenne sur plusieurs années. L'analyse de cette relation aurait été plus pertinente à l'échelle inter annuelle et du terroir, où le paysan prend en compte les données pluviométriques pour décider de l'allocation des terres.

Dans l'ensemble, on constate comme le montre la figure 9.7 que les performances des modèles de prédiction obtenus pour les zones de culture pluviale restent assez faibles dans les trois cas ($0,66 < \text{ROC} < 0,7$). On peut voir que la courbe ROC est plus proche de la diagonale, qui correspond à un modèle aléatoire. La difficulté à expliquer la distribution des zones de culture pluviales proviendrait en grande partie de sa nature. En effet cette classe d'utilisation de l'espace ne correspond pas exactement aux zones cultivées en culture pluviale. Elle inclut également les jeunes jachères constituées de savanes arbustives qui se confondent aux zones cultivées qui au moment de l'acquisition des images ne sont pas mises en culture. Les erreurs d'excédent qui traduisent ce mélange de classes ont été estimées à 42%. Ces erreurs pourraient donc introduire une composante aléatoire dans les modèles construits.

1987		1999		Changement	
Variables	β_i	Variables	β_i	Variables	β_i
β_0	-1,21378	β_0	-0,53016	β_0	-8.617
Rurpot87	+0,00020	Rurpot99	+0,00014	rainfall	+0.008
Soilpot1	+0,53400	Soilpot1	+0,32335	Soilpot1	-0.198
Soilpot2	-0,85165	Soilpot2	-1,14917	Soilpot2	-0.897
Soilpot3	-0,43085	Soilpot3	-0,14564	Soilpot3	+0.424
Soilpot5	-0,00561	Soilpot5	+0,12014	Soilpot5	+0.231
Soilpot6	+0,03860	Soilpot6	+0,15240	Soilpot6	-0.285
ROC = 0,695		ROC = 0,695		ROC = 0,665	

Tableau 9.12 : Paramètres des modèles de régression finaux de l'utilisation de l'espace « Culture pluviale ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, *** pour $0,1 < P < 0,9$.

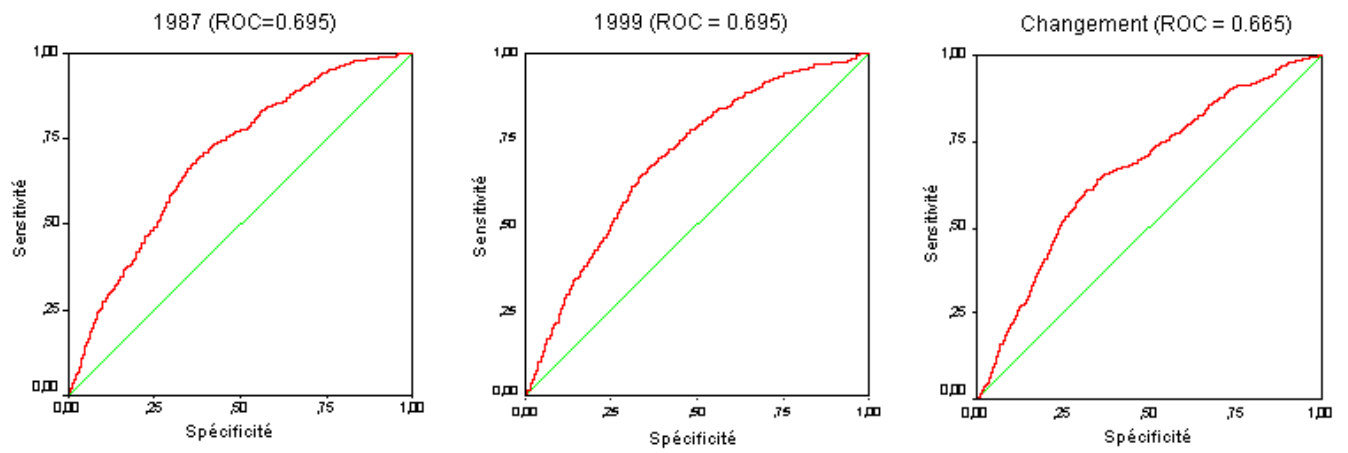


Figure 9.7 : Courbes ROC des modèles de régression finaux de l'utilisation de l'espace « culture pluviale ».

9.4.4. Les zones de sol nu

Le tableau 9.13 donne résultats de toutes les régressions simples qui sont utilisés pour la construction du modèle de distribution des zones de sol nu.

	1987		1999		Changement	
	β	ROC	β	ROC	β	ROC
Altitude	-5,1 x10 ⁻⁴	0,480**	-0,003	0,556	-0,005	0,628
Slope	-0,021	0,453**	-0,041	0,467*	-0,12755	0,546
Soilpot		0,628		0,669		0,651
Soilpot1	-0,189		-0,475		+0,553	
Soilpot2	+0,009		-0,728		-0,649	
Soilpot3	+0,406		0,283		+0,499	
Soilpot4	1,017		0,653		+0,715	
Soilpot5	-0,109		-1,051		-0,758	
Soilsorg		0,608		0,662		0,704
Soilsorg1	-1,024		-1,319		-0,549	
Soilsorg2	-0,758		-1,250		-1,352	
Soilsorg3	-,010		0,521		1,402	
Soilsorg4	-1,017		-1,001		+0,189	
Soilsorg5	-0,543		-0,927		-0,889	
Rainfall	+15x10 ⁻⁴	0,520*	-0,002	0,530	-0,004	0,571
Densite	-3,1x10 ⁻⁴	0,522*	-0,001	0,547*	-0,002	0,505
Rurpot	-1,02x10 ⁻⁴	0,527	-1x10 ⁻⁴	0,535	-1 x10 ⁻⁴	0,507
Poppot	-0,65x10 ⁻⁴	0,532	-0,98x10 ⁻⁴	0,550	-1,70 x10 ⁻⁴	0,519
Fulbé	-0,063	0,502***	0,542	0,549	+1,059	0,603
Guiziga	-0,142	0,518*	-0,288	0,534*	-0,769	0,576
Mofu	+0,279	0,518	-0,185	0,529*	-0,240	0,531*
Distriver	-0,17x10 ⁻⁴	0,500*	-0,2x10 ⁻⁴	0,509***	+0,22x10 ⁻⁴	0,502*
Distroad	+1,3x10 ⁻⁴	0,547	+0,18 x10 ⁻⁴	0,515***	-3,86x10 ⁻⁴	0,614
Travelc	+0,2x10 ⁻⁴	0,526	+0,1x10 ⁻⁴	0,523*	-0,3x10 ⁻⁴	0,550
Travelu	+0,1x10 ⁻⁴	0,510*	+0,1x10 ⁻⁴	0,512***	-2,2x10 ⁻⁶	0,481***
Travelm	+0,2x10 ⁻⁴	0,542	+0,3x10 ⁻⁴	0,550	-0,26x10 ⁻⁴	0,520
Livestock	-0,001	0,515***	-0,001	0,512***	0,005	0,534*

Tableau 9.13 : Paramètres des régressions simples pour l'utilisation « sol nu ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, * pour $0,1 < P < 0,9$.**

Le tableau 9.13 montre que presque toutes les corrélations des variables avec l'utilisation de l'espace « sol nu » ne sont pas significatives. C'est le cas des variables de relief (*altitude*, *slope*), de pluviométrie (*rainfall*) et des variables géoéconomiques (*distriver*, *distroad*, *travelc*, *travelu*, *travelm* et *livestock*). On aurait espéré avoir une corrélation importante entre les pentes et la distribution des sols nus. Cette tendance a d'ailleurs été observée sur les images satellites de cette région mais principalement dans la zone des piémonts. Toutefois, il apparaît que les pentes sont assez faibles sur l'ensemble de la zone qui est une région de plaine. Les variables ethniques (*fulbe*, *guiziga* et *mofu*) et de population (*density*, *rurpot* et *poppot*) affichent également des corrélations très faiblement significatives et des pouvoirs explicatifs assez bas (valeur ROC < 0,6).

Seules les variables de type de sol (*soilsorg* et *soilpot*) ont un pouvoir explicatif supérieur à 0,6 et ont été sélectionnées pour la régression conditionnelle pas à pas. En suivant une approche purement inductive, on aurait un modèle de régression final avec une seule variable *soilpot* dont les performances sont celles du tableau 9.14 (ROC = 0,651). En formulant l'hypothèse que la pression humaine peut expliquer le processus de défrichement des savanes boisées et par conséquent la dénudation des sols, les variables *rurpot* et *poppot* ont été également sélectionnées pour la régression conditionnelle.

Pour la carte de 1987, ces quatre variables sont insérées dans la procédure de régression conditionnelle. La variable *rurpot* et *soilsorg* ont été exclues de l'équation du modèle à l'issue de cette procédure. Le modèle de régression final obtenu affiche un pouvoir explicatif évalué à 0,640. Pour la carte de 1999, la même combinaison de variables est insérée dans la procédure de régression conditionnelle. Les quatre variables sont maintenues dans l'équation du modèle obtenu qui affiche une capacité de discrimination égale à 0,697. La corrélation avec la variable *rurpot* n'est pas significative bien que celle-ci soit maintenue dans l'équation. Ainsi, en maintenant uniquement les variables *soilpot* et *poppot*, on obtient un modèle final dont la capacité de discrimination est estimée à 0,687. Contrairement aux attentes, le signe du coefficient de la variable *poppot* indique que la probabilité que l'utilisation de l'espace soit «sol nu » est faible dans les zones où la pression humaine est forte. La cartographie et les observations de terrain sur cette région autour de Maroua, confirment cette tendance dans la mesure où les zones avoisinant les zones d'habitation sont le plus souvent suffisamment arborées. D'autre part, les parcs arborés entretenus par les populations sont visibles dans le paysage au centre de la plaine du Diamaré. Cet exemple traduit les signes d'intensification observés sur certains sites.

Pour la situation des changements entre 1987 et 1999, seules les variables *altitude* et les variables de type de sol (*soilsorg* et *soilpot*) affichent un pouvoir explicatif supérieur à 0,6. La relation avec la variable *altitude* ne traduisant pas un processus pertinent, celle-ci n'a pas été sélectionnée. Une approche strictement inductive conduirait à un modèle avec la seule variable *soilsorg* qui affichent un pouvoir explicatif de 0,704. Les changements observés pour l'utilisation de l'espace « sol nu » sont relativement importants et correspondent principalement à des récupérations de sols dégradées pour les cultures pluviales ou les cultures de sorgho de contre saison. Dans certains cas, ce sont des reboisements. Chacune de ces conversions de l'utilisation de l'espace nécessite des investissements financiers, en terme de travail, ou en terme de technique de préparation de la terre. Les variables *distroad* et *travelc* qui traduisent l'accessibilité et dans une moindre mesure la facilité à y investir ont été introduites dans la régression conditionnelle avec les variables de type de sol (*soilsorg* et *soilpot*). La variable *travelc* est exclue de l'équation du modèle à l'issue de la régression conditionnelle et le modèle obtenu à une capacité de

détection estimée à 0,722. En gardant uniquement les variables *soilsorg*, *distroad* et *travelc* on obtient un modèle final dont la capacité de détection des changements est estimée à 0,718. Une interprétation plausible du modèle obtenu c'est que le processus de récupération des terres dégradées pour l'agriculture s'effectue plus probablement dans des zones éloignées du centre urbain et qui sont accessibles (proche des routes) et donc moins contraignants pour y investir.

Parmi toutes les utilisations de l'espace, le modèle de distribution des sols nus est celui qui a la plus faible capacité de discrimination. Contrairement au cas des zones de culture pluviale dont la cartographie affiche des erreurs d'excédents importants (48%), la cartographie des sols nus affiche des erreurs d'excédents relativement moins importants (35%). La difficulté de prédiction est certes liée aux erreurs de cartographie mais une importante partie est liée à l'absence d'une variable explicative pertinente parmi les variables spatiales utilisées. A titre d'exemple, les vérités de terrain ont montré que la plupart des sols nus sont des terres dégradées et marginales. Une explication pertinente de la distribution des changements de cette forme d'occupation du sol serait liée aux pratiques culturales (extensives ou intensives). La diversité des types des conversions et des causes possibles explique d'une part la difficulté à expliquer les changements de cette forme d'occupation du sol.

1987		1999		Changement	
Variables	β_i	Variables	β_i	Variables	β_i
β_0	-2,03861	β_0	-2,15784	β_0	-3,32732
Poppot87	-0,00006	Poppot99	-0,00010	distroad	-0,00035
Soilpot 1	-0,11986	Soilpot 1	-0,34942	travelc	+4,4 x 10 ⁻⁶
Soilpot 2	-0,00247	Soilpot 2	-0,74480	Soilsorg 1	-0,47572
Soilpot 3	+0,37225	Soilpot 3	+0,22891	Soilsorg 2	-1,25301
Soilpot 5	+1,02784	Soilpot 5	+0,68880	Soilsorg 3	+1,43771
Soilpot 6	-0,02609	Soilpot 6	-0,88916	Soilsorg 5	+0,08908
				Soilsorg 6	-0,85209
ROC = 0,640		ROC = 0,687		ROC = 0,718	

Tableau 9.14 : Paramètres des modèles de régression finaux de l'utilisation de l'espace « sol nu ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, * pour $0,1 < P < 0,9$.**

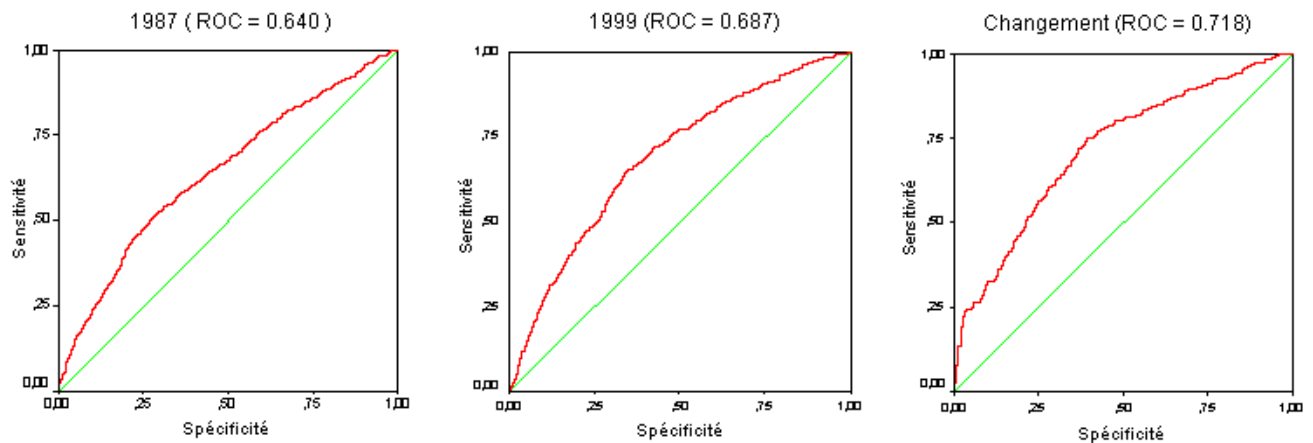


Figure 9.8 : Courbes ROC des modèles de régression finaux de l'utilisation de l'espace « sol nu ».

9.4.5. Les zones de culture du sorgho de contre saison

Le tableau 9.15 donne les résultats de toutes les régressions simples, qui sont utilisés pour construire le modèle de prédiction de l'utilisation de l'espace « sorgho repiqué ».

	1987		1999		Changement	
	β	ROC	β	ROC	β	ROC
Altitude	-53,9x10 ⁻⁴	0,561	-0,009	0,622	-70,4 x10 ⁻⁴	0,596
Slope	-0,464	0,646	-0,751	0,686	-0,609	0,642
Soilpot		0,713		0,749		0,701
Soilpot1	+0,546		+0,888		+0,535	
Soilpot2	+1,362		+2,148		+1,744	
Soilpot3	+0,166		+0,962		+1,161	
Soilpot4	-0,327		+0,192		+0,329	
Soilpot5	-1,861		-1,553		-1,366	
Soilsorg		0,735		0,777		0,715
Soilsorg1	+1,690		+2,325		1,179	
Soilsorg2	+0,861		+0,988		+0,794	
Soilsorg3	-1,303		-0,650		-0,182	
Soilsorg4	+0,344		+0,151		-0,263	
Soilsorg5	-1,487		-2,037		-2,063	
Rainfall	+0,9x10 ⁻⁴	0,506***	-0,003	0,537	-5,8x10 ⁻⁴	0,509*
Densite	-7,1x10 ⁻⁴	0,538	-4,8x10 ⁻⁴	0,542	-3,8x10 ⁻⁴	0,560
Rurpot	-3,1x10 ⁻⁴	0,564	-1,9 x10 ⁻⁴	0,558	-2,3x10 ⁻⁴	0,573
Poppot	-0,7x10 ⁻⁴	0,553	-0,3 x10 ⁻⁴	0,541	-0,2x10 ⁻⁴	0,560
Fulbé	+0,376	0,547	+0,620	0,571	+0,323	0,537
Guiziga	+0,435	0,558	+0,279	0,543	+0,408	0,551
Mofu	-1,016	0,555	-1,387	0,585	-1,379	0,588
Distriver	-0,2x10 ⁻⁴	0,501*	-0,2x10 ⁻⁴	0,496*	+0,3x10 ⁻⁴	0,545
Distroad	+0,7x10 ⁻⁴	0,530	+0,87x10 ⁻⁴	0,543	+1,2x10 ⁻⁴	0,558
Travelc	-0,2x10 ⁻⁴	0,539	-0,2x10 ⁻⁴	0,539	+0,1x10 ⁻⁴	0,523
Travelu	+0,2x10 ⁻⁴	0,527	+0,3x10 ⁻⁴	0,545	-2,57827	0,585
Travelm	-0,5x10 ⁻⁴	0,576	-0,3x10 ⁻⁴	0,547	+0,08x10 ⁻⁴	0,525
Livestock	+0,028	0,643	+0,027	0,632	+0,008	0,545

Tableau 9.15 : Paramètres des régressions simples pour l'utilisation de l'espace « sorgho de contre saison ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, * pour $0,1 < P < 0,9$.**

On remarque dans le tableau 9.15 que 2 variables affichent des corrélations non significatives. Il s'agit de *rainfall* et *distriver* qui n'ont donc pas été sélectionnées pour la régression conditionnelle. Les variables de population (*density*, *rurpot* et *poppot*), d'ethnicité (*fulbé*, *guiziga*, *mofu*) et les variables géoéconomiques (*distroad*, *travelc*, *travelu*, *travelm*) affichent une corrélation significative mais leurs pouvoirs explicatifs sont très bas ($ROC < 0,6$).

Comme nous l'avons montré au chapitre 7, les variables de type de sol (*soilpot* et *soilsorg*) sont celles qui expliquent le mieux la distribution des zones de sorgho de contre saison ($ROC > 0,7$). La variable *livestock* affiche un pouvoir explicatif relativement plus bas ($0,6 < ROC < 0,7$). Toutefois, une hypothèse a été formulée au chapitre 7 sur la tendance d'intégration entre la culture du sorgho de contre saison et le développement de l'élevage. Cette variable a donc été sélectionnée pour la procédure de régression conditionnelle. Contrairement à la variable *altitude* dont le pouvoir explicatif est assez bas, la variable *slope* affiche une corrélation significative et un pouvoir explicatif relativement plus élevé ($0,6 < ROC < 0,7$). Ce résultat traduit le fait que l'altitude en soit n'est pas un facteur contraignant pour la mise en culture du sorgho repiqué, mais plutôt la pente. Une approche inductive basée uniquement sur l'exploration des données aboutit à un modèle avec les variables de type de sol (*soilsorg* et *soilpot*), de relief (*slope*) et la variable *livestock*.

Pour la situation de l'année 1987, toutes ces quatre variables sont maintenues dans l'équation du modèle à l'issue de la régression conditionnelle. Le modèle de régression ainsi construit fournit une performance évaluée à 0,770. Les variables *rurpot* et *travelm* qui ont des pouvoirs explicatifs avoisinant 0,6, sont incluses dans le modèle pour prendre en compte l'influence du marché et de la pression humaine. Dans ce cas, on obtient un modèle qui a une plus grande capacité de discrimination des zones de sorgho de contre saison, soit $ROC=0,776$.

Pour la situation de 1999, On obtient des résultats similaires. En suivant l'approche purement inductive, on obtient un modèle qui a une capacité de discrimination évaluée à 0,807. En insérant les variables *travelm* et *rurpot*, la capacité de discrimination du modèle obtenu est plus importante, soit $ROC = 0,810$. Comme le montre les courbes ROC de la figure 9.9, les deux modèles produits fournissent une assez bonne prédiction de la distribution spatiale des zones de sorgho de contre saison. On peut remarquer dans le tableau 9.16 que les signes des coefficients de toutes les variables incluses dans ces modèles coïncident entre les deux dates et fournissent des explications pertinentes. En effet, la probabilité d'affectation d'un espace à la culture du sorgho de contre saison diminue avec la pente et la distance au marché. Elle augmente avec la densité du cheptel bovin. Ce qui confirme la complémentarité entre l'élevage transhumant et la culture du sorgho repiqué qui a été expliquée amplement au chapitre 7.

Pour la distribution spatiale des changements entre 1987 et 1999, les variables sélectionnées en suivant une approche inductive sont : *slope*, *soilsorg* et *soilpot*. Le modèle obtenu a une capacité de détection des changements estimée à 0,735. La variable *rurpot* qui affiche un pouvoir explicatif proche de 0,6 a été incluse dans le modèle pour prendre en compte l'influence de la pression humaine sur les changements observés. Le modèle obtenu affiche une capacité de détection des changements évaluée à 0,733 mais, affiche une corrélation non significative avec la variable *rurpot*. Ce résultat s'explique par l'organisation de l'utilisation de l'espace autour des zones d'habitation dans la région. Les zones de culture de sorgho repiqué sont en général localisées dans la dernière auréole avant les zones de brousse c'est à dire très éloignées du centre du village. Par contre la variable *rurpot* affiche les plus grandes valeurs au centre du village et prend des valeurs décroissantes lorsqu'on se déplace du centre du village vers les zones de

brousse. La variable *rupot* a donc été exclue de l'équation du modèle final qui affiche la même capacité de détection des changements (ROC = 0,733).

Les changements observés pour l'utilisation de l'espace « sorgho de contre saison » correspondent principalement à des espaces transformés en brousse (longues jachères) ou à des conversions en culture pluviale. Une faible proportion est convertie en culture maraîchère ou en sol nu. Les conversions en culture pluviale ou en culture maraîchère se font sur les sols intermédiaires entre les vertisols et les sols d'apport alluvial. Les longues jachères observées concernent le plus souvent les terres les plus aptes à la culture du sorgho de contre saison. Ceci est confirmé par la corrélation positive avec la variable *soilsorg* qui affiche un pouvoir explicatif élevé (ROC > 0,7). Ces jachères surviennent le plus souvent après plusieurs années d'exploitation successives, qui pourrait traduire ainsi une demande de plus en plus importante de la population. En plus des conditions du milieu (type de sol et relief), les changements observés sont fortement dictés par les besoins de production plus importants qui se traduisent par les pratiques paysannes liées à l'amélioration de la fertilité des sols, mais qui ne peuvent être capturées dans ce type de modèle.

1987		1999		Changement	
Variables	β_i	Variables	β_i	Variables	β_i
β_0	-2,01523	β_0	-1,56436	β_0	-1,737
Slope	-0,21279	Slope	-0,42343	Slope	-0,363
Livestock	+0,01968	Livestock	+0,01857	Soilsorg1	+0,939
Rurpot87	-0,00031	Travelm	-0,00002	Soilsorg 2	+0,628
Travelm	-0,00004	Rurpot99	-0,00009	Soilsorg 3	-0,345
Soilsorg1	+1,36185	Soilsorg1	+1,93026	Soilsorg 5	-0,480
Soilsorg2	+0,84024	Soilsorg2	+0,88849	Soilsorg 6	-1,563
Soilsorg3	-1,41260	Soilsorg3	-0,82629		
Soilsorg5	+0,21811	Soilsorg5	-0,11627		
Soilsorg6	-0,92469	Soilsorg6	-1,36438		
ROC = 0,776		ROC = 0,810		ROC = 0,733	

Tableau 9.16 : Paramètres des modèles de régression finaux de l'utilisation de l'espace « sorgho de contre saison ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, *** pour $0,1 < P < 0,9$.

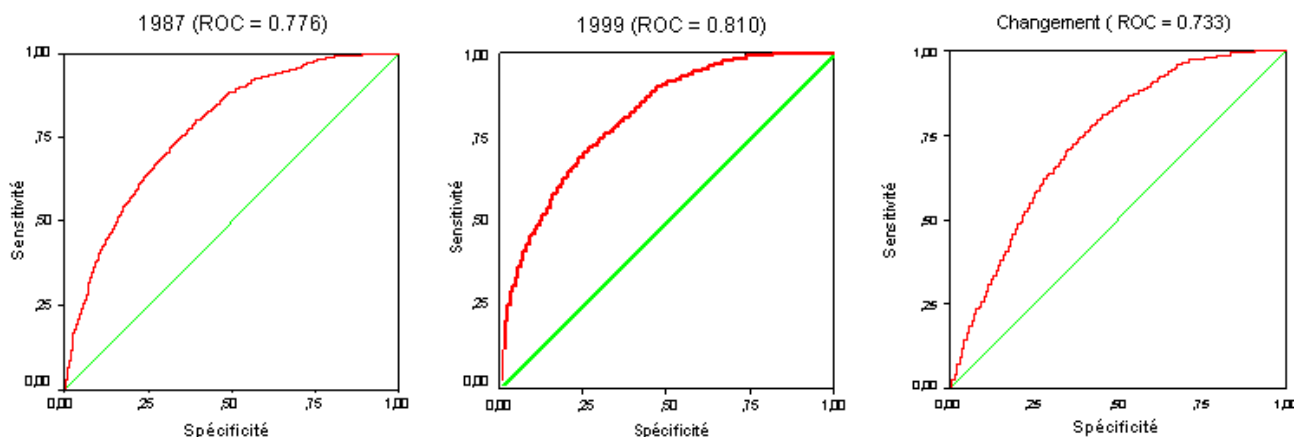


Figure 9.9 : Courbes ROC des modèles de régression finaux de l'utilisation de l'espace « sorgho de contre saison ».

9.4.6. Les zones de brousse ou savanes boisées

Le tableau 9.17 donne les résultats notamment, les paramètres statistiques de toutes les régressions simples qui sont utilisées pour construire les modèles de prédiction de l'utilisation de l'espace « brousse ».

	1987		1999		Changement 1987 - 199	
	β	ROC	β	ROC	β	ROC
Altitude	-0,005	0,584	-0,005	0,575	+1,3x10 ⁻⁴	0,587***
Slope	-0,237	0,574	-0,294	0,564	-0,178	0,536
Soilpot		0,635		0,630		0,635
Soilpot1	-1,097		-1,477		-0,140	
Soilpot2	-0,122		-0,395		+0,869	
Soilpot3	0,320		-0,323		+0,382	
Soilpot4	-0,497		-0,621		+0,200	
Soilpot5	-1,790		-2,562		-1,364	
Soilsorg		0,624		0,650		0,657
Soilsorg1	-0,405		-0,996		-0,528	
Soilsorg2	+0,363		+0,701		+0,983	
Soilsorg3	+0,296		+0,693		-0,019	
Soilsorg4	-0,905		-0,864		-0,314	
Soilsorg5	-0,603		-0,093		-0,432	
Rainfall	+0,002	0,530	+0,001	0,530*	+0,012	0,668
Densite	-3,9x10 ⁻⁴	0,648	-0,011	0,680	-0,003	0,593
Rurpot	-6,2x10 ⁻⁴	0,669	-0,001	0,726	-0,001	0,658
Poppot	-4,5 x10 ⁻⁴	0,684	-0,001	0,736	-0,001	0,666
Fulbé	+0,081	0,493***	+0,171	0,502*	-0,708	0,565
Guiziga	+0,486	0,531	+0,610	0,546	+0,958	0,594
Mofu	-1,297	0,619	-2,011	0,638	-0,644	0,558
Distriver	+1,6x10 ⁻⁴	0,644	+1,5x10 ⁻⁴	0,646	+0,7x10 ⁻⁴	0,575
Distroad	+1,98x10 ⁻⁴	0,579	+3,1x10 ⁻⁴	0,620	+3,14x10 ⁻⁴	0,610
Travelc	+1,03x10 ⁻⁴	0,671	+1,3x10 ⁻⁴	0,707	+0,76x10 ⁻⁴	0,643
Travelu	+0,97x10 ⁻⁴	0,632	+1,3x10 ⁻⁴	0,675	+1,05x10 ⁻⁴	0,637
Travelm	+1x10 ⁻⁴	0,655	+1,2x10 ⁻⁴	0,694	+0,3x10 ⁻⁴	0,569
Livestock	-0,014	0,565	-0,018	0,577	-0,013	0,561

Tableau 9.17 : Paramètres des régressions simples pour l'utilisation de l'espace « brousse ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, * pour $0,1 < P < 0,9$.**

On remarque dans le tableau 9.17 que les corrélations avec les variables d'altitude (*altitude*), de pluviométrie (*rainfall*) ne sont pas significatives respectivement pour la situation des changements et la situation de 1999. La corrélation avec la variable d'ethnicité (*fulbé*) est également non significative. Presque toutes les autres variables affichent une corrélation significative mais leur pouvoir explicatif reste très faible ($ROC < 0,6$). Il s'agit des variables d'ethnicité (*guiziga mofu*), de la variable d'altitude (*slope*) et la variable *livestock*. Par contre, les variables *distriver* et *travelm* qui ont un pouvoir explicatif relativement plus élevé ($0,6 < ROC < 0,7$) n'ont pas été sélectionnées parce qu'elles ne traduisent pas un processus pertinent ou une hypothèse que l'on veut tester.

Les variables géoéconomiques (*travelc*, *travelu*) et les variables de population (*poppot* et *rurpot*) affichent les meilleurs pouvoirs explicatifs ($ROC > 0,7$). Ces variables ont été sélectionnées pour la régression conditionnelle pas à pas puisque les corrélations avec ces variables traduisent la pression de la population et la proximité de la ville avec leurs besoins en bois et aliments qui sont les principales incitations aux défrichements des savanes boisées dans la région. Les variables de type de sol (*soilsorg* et *soilpot*), les variables *density* et *distroad* qui ont un pouvoir explicatif compris entre 0,6 et 0,7 ont été également sélectionnées. On a vu aux chapitres 7 et 8 à partir de l'analyse des images satellites confirmée sur le terrain que la plupart des défrichements des savanes boisées se font principalement pour l'extension agricole. On peut dans ce cas s'attendre à ce que les types de sol expliquent la distribution spatiale des zones de brousse et des changements observés. On pourrait également envisager que les zones de brousse restantes soient localisées à une grande distance des principaux axes routiers.

Pour la situation de 1987, les 8 variables suivantes ont été incluses dans la procédure de régression conditionnelle : *travelc*, *travelu*, *poppot*, *rurpot*, *soilpot*, *soilsorg*, *density* et *distroad*. Les variables *density*, *rurpot* et *soilsorg* ont été exclues de l'équation du modèle à l'issue de cette procédure. Le modèle final obtenu a une capacité de discrimination évaluée à 0,720. Pour la situation de 1999, la même combinaison de variables que celle de 1987 a été sélectionnée pour la procédure de régression conditionnelle et les mêmes variables sont maintenues dans l'équation du modèle final. Le modèle de prédiction dans ce cas affiche une capacité de discrimination plus élevée ($ROC = 0,778$). Les signes des coefficients logistiques, coïncident entre les deux dates mais seules les valeurs diffèrent. Les signes des coefficients des variables *distroad* et *travelc* indiquent que la probabilité pour qu'un espace donné soit encore boisé augmente avec la distance aux routes et au principal centre urbain (tableau 9.18). Le coefficient de la variable *travelu* est par contre difficile à interpréter. Puisque chaque centre urbain définit un front de défrichement qui part de son centre, on aurait espéré avoir le même sens de corrélation qu'avec la variable *travelc*.

Pour la situation des changements entre 1987 et 1999, la même combinaison de variables que celles des deux cas précédents a été insérée dans la procédure de régression conditionnelle. On constate qu'en plus des variables rejetées dans les deux premiers cas, la variable *travelu* a été également rejetée de l'équation du modèle. La capacité de discrimination du modèle obtenu est évaluée à $ROC = 0,707$. Les changements observés pour l'utilisation de l'espace « brousse » correspondent aux défrichements de la savane boisée. Les corrélations positives avec *distroad* et *travelc* indiquent que la probabilité qu'une zone ait subi un défrichement au cours de la période considérée est plus grande dans des zones difficilement accessibles à partir des routes ou de la ville. En effet, l'analyse diachronique de l'occupation du sol effectuée au chapitre 8 a montré que les défrichements se sont effectués dans le temps suivant un gradient partant du

centre de la ville aux zones rurales environnantes. Ces facteurs seraient donc plus explicatifs de la distribution spatiale des brousses restantes que des changements observés. Les observations de terrain ont confirmé l'influence du type de sol sur la présence de reliques de zones boisées dans la région. En effet, les zones encore boisées correspondent le plus souvent à des situations où les sols présentent des très faibles potentialités agricoles et où les essences d'arbres présentes ne sont pas appréciées pour le bois de feu.

1987		1999		Changement	
Variabes	β_i	Variabes	β_i	Variabes	β_i
β_0	-1,12485	β_0	-1,57875	β_0	-4,19409
Distroad	+0,00005	Travelc	+0,00010	Distroad	+0,00021
Poppot87	-0,00023	Travelu	-0,00003	Travelc	+0,00003
Travelc	+0,00010	Poppot99	-0,00055	Poppot99	-0,00022
Travelu	-0,00004	Distroad	+0,00011	Soilpot1	+0,188
Soilpot1	-0,89519	Soilpot1	-1,12442	Soilpot2	+0,77
Soilpot2	-0,27048	Soilpot2	-0,63779	Soilpot3	+0,306
Soilpot3	+0,08777	Soilpot3	-0,68003	Soilpot5	+0,284
Soilpot5	-0,50455	Soilpot5	-0,51355	Soilpot6	-1,023
Soilpot6	-1,57547	Soilpot6	-2,01718		
ROC= 0,720		ROC= 0,778		ROC= 0,707	

Tableau 9.18 : Paramètres des modèles de régression finaux de l'utilisation de l'espace « brousse ». Toutes les variables sont significatives avec un niveau $P < 0,01$ dans le cas contraire les symboles * pour $0,01 < P < 0,05$, ** pour $0,05 < P < 0,1$, * pour $0,1 < P < 0,9$.**

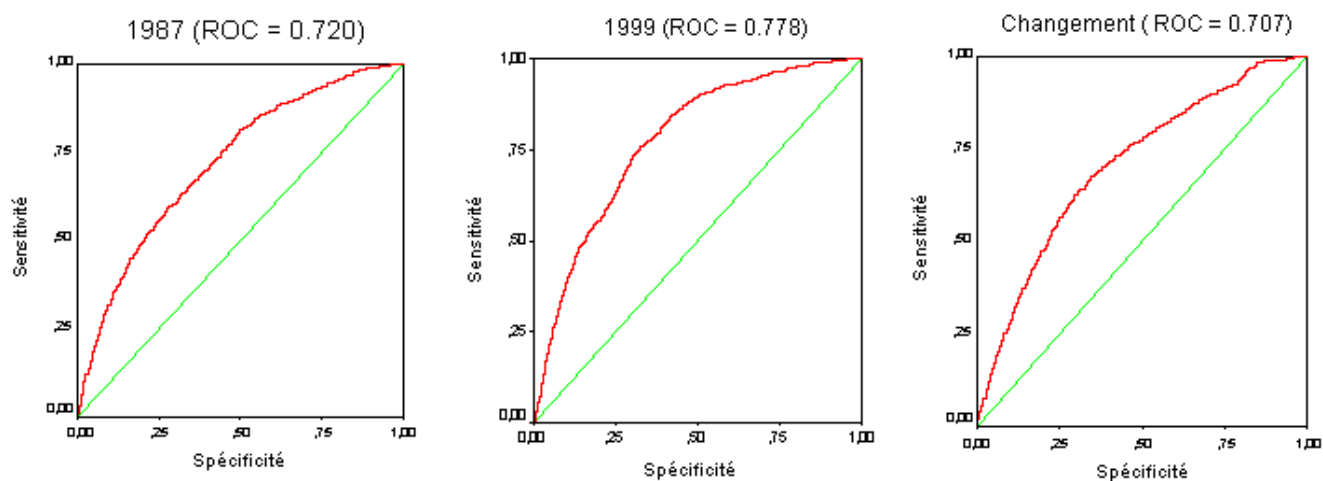


Figure 9.10 : Courbes ROC des modèles de régression finaux de l'utilisation de l'espace « brousse ».

9.5. Synthèse et discussions

L'exploration des facteurs déterminants a permis de construire un modèle de prédiction pour chaque utilisation de l'espace dans les trois situations considérées : 1987, 1999 et les changements entre les deux dates. Le tableau 9.19 résume les combinaisons des variables explicatives et les capacités de discrimination de tous les modèles de prédiction obtenus.

LU	Date	Variables des modèles de prédiction					ROC	P*		
<i>Habitation</i>	1987	Poppot87	Distroad	Soilsorg		0,886	92			
	1999	Poppot99	Distroad	Soilsorg		0,887				
	Changement	Poppot99	Distroad	Distriver	Soilsorg			0,889		
<i>Maraîcher Verges</i>	1987	Distriver	Travelc	Travelu	Travelm	Rurpop87	Poppot87	Soilsorg	0,939	94
	1999	Distriver	Travelc	Travelm	Rurpop99	Poppot99	Soilsorg		0,915	
	Changement	Distriver	Travelc	Travelu	Travelm	Rurpop99	Soilsorg		0,888	
<i>Culture Pluviale</i>	1987	Rurpot87	Soilpot					0,695	52	
	1999	Rurpot99	Soilpot					0,695		
	Changement	rainfall	Soilpot					0,665		
<i>Sol nu</i>	1987	Poppot87	Soilpot					0,640	65	
	1999	Poppot99	Soilpot					0,687		
	Changement	Distroad	Travelc	Soilsorg						0,718
<i>Sorgho repiqué</i>	1987	Slope	Livestock	Rurpot87	Travelm	Soilsorg		0,776	89	
	1999	Slope	Livestock	Travelm	Rurpot99	Soilsorg		0,810		
	Changement	Slope	Soilsorg					0,733		
<i>Brousse</i>	1987	Distroad	Poppot87	Travelc	Travelu	Soilpot		0,720	85	
	1999	Travelc	Travelu	Poppot99	Distroad	Soilpot		0,778		
	Changement	Distroad	Travelc	Poppot99	Soilpot		0,707			

Tableau 9.19 : Combinaisons des variables explicatives et performances des modèles de prédiction des utilisations de l'espace. La colonne *P représente la précision d'utilisation des classes d'occupation du sol calculée au chapitre 8. Elle est d'autant plus faible que les erreurs d'excédent sont importantes.

9.5.1. Aperçu général sur les facteurs explicatifs de l'utilisation de l'espace et des changements

On peut remarquer dans le tableau 9.19 que les natures des variables qui expliquent la distribution spatiale de chaque utilisation de l'espace (structure du système) ne changent pas fondamentalement entre 1987 et 1999. Les signes des coefficients logistiques coïncident également entre les deux dates. Les changements les plus importants concernent l'importance relative des valeurs des coefficients qui traduisent l'influence des différents facteurs. Toutefois, on note quelques exemples de situations qui sont difficiles à expliquer. L'absence de la variable *travelu* dans l'équation du modèle des cultures maraîchères/verges en 1999 et le signe de son coefficient dans le modèle de la brousse sont contraires aux attentes si on considère l'hypothèse formulée sur l'influence des petits centres urbains de la région

(chapitre 5). On pourrait conclure que cette influence n'est pas encore assez marquée comme celle du principal centre urbain. Les facteurs déterminant les changements ne sont pas toujours les mêmes qui déterminent la distribution spatiale, ce qui est bien compréhensible. Mais, dans plusieurs des cas, on ne dispose pas de variables spatiales pertinentes qui traduisent les processus qui gouvernent ces changements. Ainsi, plusieurs variables identifiées comme déterminant les changements traduisent en réalité la localisation des unités qui ont subi les changements.

Le tableau 9.19 montre que presque toutes les variables interviennent significativement dans l'explication des structures d'utilisation de l'espace à l'exception des variables de pluviométrie (*rainfall*) et d'ethnicité (*fulbé, guiziga, mofu*). La variable *rainfall* apparaît uniquement dans le modèle des cultures pluviales mais nous n'avons pas pu trouver une interprétation raisonnable à cette corrélation. Cette dernière situation s'explique d'une part par la précision spatiale des données et l'échelle temporelle considérée qui ne permettent pas de capturer les phénomènes étudiés. De même, on a observé des corrélations avec certaines variables d'ethnicité (*fulbé* avec les cultures maraîchères et *mofu* pour les cultures pluviales) qui ne pouvaient pas être associée à des mécanismes ou stratégies connues. Les modèles proposés ne permettent pas de capturer facilement la manière (pas nécessairement linéaire) dont ce type de variable influence l'utilisation de l'espace.

Les potentialités des sols et les variables démographiques apparaissent comme des facteurs intervenant dans l'explication de la structure et des changements de toutes les utilisations de l'espace. L'indice de population potentielle par exemple apparaît dans l'explication de la distribution de presque toutes les utilisations de l'espace sauf pour les cultures maraîchères/vergers qui obéissent plus à une logique économique. Les résultats de la procédure d'exploration ont confirmé la très forte influence du centre urbain sur l'évolution des cultures maraîchères et du processus de défrichement des savanes boisées. La pression sur l'espace ou la disponibilité des terres évaluées à partir d'un indice de population potentielle explique mieux la structure et les changements d'utilisation de l'espace que la densité de la population.

9.5.2. Modèles de prédiction de chaque utilisation de l'espace

On remarque dans l'ensemble du tableau 9.19 que chaque utilisation de l'espace a son sous-ensemble spécifique de facteurs déterminants. Dans la plupart des cas, sauf pour quelques-uns qui ont été explicités, les variables qui interviennent dans l'explication de chaque utilisation de l'espace ont généralement le signe espéré et permettent ainsi d'aboutir à des modèles de prédiction plus réalistes et interprétables. Dans cette sous-section, les résultats sont résumés et confrontés avec les hypothèses théoriques ou les connaissances empiriques pour chaque type d'utilisation de l'espace.

Zones d'habitation

La distribution spatiale des zones d'habitation s'explique principalement par la pression humaine et s'organise autour des principaux axes routiers. Les habitations ont également une tendance à s'installer sur des terres ayant un faible taux d'argile. Les changements observés pour cette classe d'utilisation de l'espace surviennent de préférence le long des cours d'eau sous la menace des effets de l'érosion des berges. Cette dynamique peut être mieux expliquée à l'échelle du terroir.

Cultures maraîchères/Vergers

L'organisation spatiale des cultures maraîchères et des vergers est déterminée par un nombre important de variables notamment celles liées au milieu physique (proximité des cours d'eau et type de sol),

celles liées à la demande (population) et à l'accessibilité au marché. L'influence de la qualité des sols sur la distribution des cultures maraîchères se traduit dans le modèle par une corrélation positive avec les sols d'apport alluvial à texture fine (*soilsorg5*) et les sols argileux à faciès verticale (*soilsorg2*). Ainsi, le modèle obtenu confirme globalement l'idée d'une organisation de la culture maraîchère le long des cours d'eau et autour du principal centre urbain en fonction des types de sol. Pour l'année 1987, le modèle indique qu'une organisation similaire se répète à l'échelle des petits centres urbains. Toutefois, cette observation n'est pas confirmée en 1999. L'influence de la variable qui définit la distance au marché est difficile à expliquer parce que les marchés considérés sont plutôt spécialisés dans la vente du bétail et du sorgho de contre saison et non dans les produits maraîchers ou les fruits.

Cultures pluviales

Aucun facteur n'explique de façon assez évidente et individuelle la localisation des cultures pluviales. La plupart des variables affichent un pouvoir explicatif assez faible. Le modèle obtenu met en exergue l'influence significative de la pression humaine et des types de sol, confirmant ainsi l'hypothèse selon laquelle les zones de cultures pluviales constituent la première auréole autour du village en zone rurale et en priorité sur les meilleures terres. L'indice de pression potentielle de la population explique mieux cette distribution que la densité de la population.

Sols nus

Tout comme pour les cultures pluviales, aucun facteur n'explique de façon assez évidente la localisation des sols nus et le même couple de variables explicatives est obtenu dans ce cas. Compte tenu de l'impact de la dégradation des sols suite à l'érosion on aurait espéré une corrélation positive avec la pente (*slope*). De même, on aurait également espéré une corrélation avec l'importance de l'élevage qui est perçu par les acteurs de la région comme un facteur qui crée la pression sur la végétation et donc la dénudation des sols. Toutefois, aucune des corrélations avec ces variables n'est significative.

Sorgho repiqué

La distribution des zones cultivées en sorgho repiqué est déterminée par une combinaison de variables de natures diverses incluant 2 variables biophysiques, une variable démographique et 2 variables géoéconomiques. Le modèle confirme l'évidence selon laquelle l'extension spatiale est fortement contrainte par le type de sol et le niveau de pente du sol. Il met également en exergue le fait que la culture du sorgho de contre saison constitue la dernière auréole autour des zones d'habitation, s'installant ainsi sur les espaces vides où la pression humaine est faible (*rurpot*). La compatibilité de l'utilisation pastorale de l'espace (*livestock*) avec la culture du sorgho repiqué est également confirmée par ce modèle. Toutefois, la précision du MNT utilisée ne permet pas de mettre en exergue les dynamiques d'extension de la culture sur les sites en microtopographie et qui exigent l'aménagement des diguettes.

Brousse

Le modèle obtenu pour la brousse montre que sa distribution spatiale est principalement déterminée par la pression humaine et l'accessibilité. En effet, la pression humaine a conduit à des défrichements sur la plupart des meilleures terres pour l'extension agricole et la satisfaction des besoins en bois de feu. Par conséquent, les zones de brousse restantes sont de plus en plus éloignées des zones d'habitation et des routes principales. Quelques observations sur le terrain ont montré que les zones

non défrichées qui sont relativement proche du centre urbain correspondent à des terres de très faibles potentialités agricoles et dont les espèces ligneuses présentes ne sont pas prisées pour le bois de feu.

9.5.3. Performances des modèles de prédiction

Une tendance générale qui se dégage du tableau 9.19 est que les performances des modèles de prédiction des utilisations de l'espace sont d'autant plus élevées que la précision de la cartographie meilleure. Ainsi, les modèles de prédictions les plus performants sont ceux de la distribution des zones de cultures maraîchères et des zones d'habitation. Les modèles des zones de culture de sorgho repiqué et des zones de brousse sont assez performants alors que les zones de cultures pluviales et de sols nus restent difficiles à prédire. Les classes culture pluviale et sol nu affichent les valeurs les plus faibles de la précision d'utilisation, traduisant les difficultés de leur identification par cartographie (chapitre 8). Les difficultés de prédictions de ces classes s'expliquent par le fait que les mélanges de classe (erreurs d'excédent) introduisent des aléas dans le modèle qui sont difficiles à expliquer. En plus de la qualité de la cartographie, la performance du modèle de prédiction dépend également de la pertinence et de la complétude des facteurs déterminants qui ont été introduits dans le modèle. On a vu dans cette application plusieurs exemples de facteurs pertinents qui n'ont pas été inclus dans les modèles soit à cause de l'absence de données spatiales, soit à cause des difficultés à exprimer ces facteurs sous forme des variables spatiales pertinentes, etc. Il se pose également un problème d'échelle. En effet, pour certaines utilisations comme les zones d'habitation et les cultures maraîchères les superficies concernées par les changements sont assez faibles et difficiles à analyser à l'échelle régionale. Les facteurs déterminant ces changements opèrent également à l'échelle locale et nécessite dans ce cas de mettre en œuvre une méthode explicitement multi-échelle.

On remarque que les modèles de prédiction des changements sont généralement moins performants que les modèles de prédiction de l'organisation spatiale. Une raison importante de cette situation est la difficulté pour le type de modèle utilisé de capturer les mécanismes qui induisent les changements. On a vu par exemple que certaines utilisations présentent une grande diversité des formes de conversions et donc de causes possibles et de stratégies mises en place. Le type de modèle utilisé est donc en réalité statique et n'est pas adapté pour prédire les changements d'utilisation de l'espace. Ces modèles permettent de calculer des probabilités de localisation des utilisations ou des changements qui sont exploités ultérieurement pour calibrer un modèle qui simule les changements d'utilisation de l'espace en prenant en compte la compétition entre les différentes formes d'utilisation de l'espace. La comparaison des résultats de plusieurs modèles indique que le modèle le plus pertinent n'est pas nécessairement celui correspondant aux hypothèses formulées à l'avance, ni celui correspondant aux meilleures corrélations fournies par les données. Il correspond à un compromis qui tend à refléter les processus étudiés et justifie l'intérêt de l'approche d'exploration utilisée.

9.6. Conclusion et perspectives

Dans ce chapitre, une méthode d'exploration spatiale a été appliquée sur un jeu de données extrait de la base de données de SMALL Savannah pour analyser les facteurs qui expliquent la structure et les changements d'utilisation de l'espace dans la région de Maroua, Extrême Nord du Cameroun. La méthode combine une approche inductive basée sur l'exploration des données avec des aspects d'un raisonnement déductif basé sur les connaissances préalables sur la zone d'étude et les théories de changement d'utilisation de l'espace. Un ensemble de variables spatiales a été calculé pour représenter l'utilisation de l'espace et traduire les facteurs qui influencent potentiellement sa structure et ses

changements. Un modèle de prédiction a été construit pour chacune utilisation de l'espace en distinguant la situation de 1987, celle de 1999 et des changements entre les deux dates. Les modèles ainsi développés fournissent une meilleure caractérisation (quantitative) des relations entre l'utilisation de l'espace et un ensemble de facteurs d'ordre biophysique, sociodémographique et géoéconomique. Les résultats confirment la très forte influence du centre urbain et révèlent celle des potentialités des terres agricoles sur les formes d'utilisation de l'espace. La pression sur l'espace évaluée à partir d'un indice de population potentielle explique mieux les structures et les changements que la densité de la population.

Cette application met en exergue l'ensemble des situations envisagées par la procédure de sélection des facteurs potentiellement déterminants introduite au chapitre 3. Des hypothèses émises sur l'influence de certains facteurs sont confirmés alors que certaines sont rejetées et de nouvelles relations pertinentes sont identifiées et quantifiées.

Sur le plan méthodologique, les performances des modèles construits dépendent de deux facteurs : l'approche d'exploration et la nature des données. La comparaison des performances de prédiction des modèles obtenus avec celles des modèles issus d'une approche purement déductive ou celle essentiellement basée sur les meilleures statistiques fournies par les données, justifie l'intérêt et la pertinence de la méthode proposée pour explorer la complexité du système d'utilisation de l'espace. Les résultats obtenus montrent qu'en combinant une approche inductive avec des aspects de raisonnement déductif, on arrive à mieux expliciter les interactions entre ces facteurs et les changements d'utilisation de l'espace. En conclusion les études portant sur l'utilisation de l'espace devraient adopter de plus en plus des approches d'analyse plus déductives et intégrées qui prennent en compte les processus réels et la multiplicité des interactions entre l'utilisation de l'espace et un ensemble de facteurs déterminants. Les principales limites de ce type d'approche sont liées à la disponibilité et à la qualité des données. Les performances des modèles de prédiction dépendent de la qualité de la cartographie, de la pertinence des facteurs insérés dans les modèles et de l'échelle à laquelle opère le processus. Certains processus bien identifiés sont difficiles à modéliser sous forme des variables spatiales explicites. Quelques fois, la modélisation peut être simple mais le problème se pose au niveau des moyens de collecte des données spatiales explicites. De plus toutes les données ne sont pas collectées dans le même contexte ni pour les mêmes objectifs, ce qui explique la variabilité des échelles de représentation qui ont une influence certaine sur les résultats de régression et affectent les performances de prédiction des modèles. En conclusion le développement des bases de données spatiales explicites contenant des données socio-démographiques est requis pour les études sur les changements d'utilisation de l'espace.

Les performances des modèles peuvent être également améliorées en explorant la contribution d'autres échelles spatiales qui ont été jugées pertinentes. La même analyse effectuée dans ce chapitre peut être par exemple répétée aux résolutions de 500 m et 750 m. D'autre part, nous avons montré au chapitre 8, l'intérêt d'une analyse multi-échelle prenant en compte les zones agro écologiques pour lesquelles des modèles plus spécifiques peuvent être construits. L'exploration des facteurs déterminants décrite dans ce chapitre fournit uniquement des éléments de caractérisation de la structure du système d'utilisation de l'espace. Le type de modèle utilisé est donc en réalité statique et n'est pas adapté pour prédire les changements d'utilisation de l'espace. Toutefois, les informations quantitatives dérivées sont utiles pour une modélisation dynamique prenant en compte les mécanismes de changements. C'est dans cette perspective que les résultats de l'analyse spatiale et quantitative effectuée dans ce chapitre seront utilisés au chapitre suivant pour construire un modèle intégré de simulation des changements et de la compétition entre les différentes formes d'utilisation de l'espace.