Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model

Suzanne Serneels*, Eric F. Lambin

Department of Geography, University of Louvain, Louvain, Belgium

Abstract

This study attempts to identify how much understanding of the driving forces of land-use changes can be gained through a spatial, statistical analysis. Hereto, spatial, statistical models of the proximate causes of different processes of land-use change in the Mara Ecosystem (Kenya) were developed, taking into account the spatial variability of the land-use change processes. The descriptive spatial models developed here suggest some important factors driving the land-use changes that can be related to some well-established theoretical frameworks. The explanatory variables of the spatial model of mechanised agriculture suggest a von Thünen-like model, where conversion to agriculture is controlled by the distance to the market, as a proxy for transportation costs, and agro-climatic potential. Expansion of smallholder agriculture and settlements is also controlled by land rent, defined, in this case, by proximity to permanent water, land suitability, location near a tourism market, and vicinity to villages to gain access to social services (e.g. health clinics, schools, local markets). This difference in perception of land rent reflects the widely different social and economic activities and objectives of smallholders versus the large entrepreneurs involved in mechanised farming. Spatial heterogeneity as well as the variability in time of land-use change processes affect our ability to use regression models for wide ranging extrapolations. The models allow evaluating the impact of changes in driving forces that are well represented by proximate causes of land-use change. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Multiple logistic regression; Land-use change; Proximate causes; Kenya

1. Introduction

Spatially explicit modelling of land-use changes is an important technique for describing processes of change in quantitative terms and for testing our understanding of these processes. Many different modelling approaches have been adopted in the study of land-use/land-cover change. Models of land-use change can be aimed at predicting the spatial pattern of changes — addressing the question "*where* are land-use changes taking place?" — or the rates of change — addressing the question "at what *rate* are

land-cover changes likely to progress?" These two questions have been referred to as the *location* issue versus the *quantity* issue (Pontius and Schneider, 2001). The spatial prediction of land-use changes only requires an understanding of the proximate causes of the changes. The projection of future rates of land-use change is a more difficult task as it requires a good understanding of the underlying driving forces of the change (Riebsame et al., 1994). Driving forces are sometimes remote in space or in time from the observed changes, and often involve macro-economic transformations and policy changes that are difficult to anticipate.

Several studies have achieved a good projection of likely patterns of land-use change, based on multivariate models representing the interactions between the

* Corresponding author. Tel.: +32-10-472992;
fax: +32-10-472877.
E-mail address: serneels@geog.ucl.ac.be (S. Serneels).

natural and cultural landscape variables that are controlling these changes (e.g. Chomitz and Gray, 1996; Mertens and Lambin, 2000). Such spatial statistical models attempt to identify explicitly the proximate causes of land-cover changes using multivariate analyses of possible exogenous contributions to empirically-derived rates of change. Multiple linear regression techniques are generally used for this purpose. A difficulty in applying statistical models to land-use change studies lies in their inability to deal with spatial variability in the processes of land-use change. In most regions, there is a high geographic variability in land-cover types, biophysical and socio-economic drivers of land-use changes, or institutions (including policies). This spatial heterogeneity leads to variability in the causes and processes of land-use changes. Different model parameterisations are thus required for different land-use change processes (Lambin, 1994). This poses the important problem of the optimal definition of the spatial entities or regions for which a given statistical land-use change model should be calibrated. On the one hand, the spatial aggregation of a diversity of land-use change processes in a single model, calibrated over a large area, is likely to weaken the predictive ability of this model. On the other hand, subdividing a region in many small homogeneous spatial entities and calibrating different models for every entity is not compatible with the quest for generality behind any modelling activity (Mertens and Lambin, 1997). This issue will be addressed in this study through a segmentation of the territory in areas that are homogeneous in terms of land-use change processes.

However, the main theme of this study concerns the depth of understanding of land-use change processes that one can gain through descriptive models of spatial relationships of land-use changes. Do spatial, statistical models of land-use change have a value beyond the prediction of the spatial patterns of change? Some authors have argued that one can infer information on the process and dynamics of a system on the basis of observations of patterns representing the configuration of this system (Liverman et al., 1998). For example, certain categories of land-use changes tend to fragment the landscape (e.g. expansion of smallholder farming, small-scale logging, overgrazing around deep wells). Other land-use changes increase landscape homogeneity (e.g. mechanised cultivation or ranching

over large areas). Some authors have shown that there is a good correlation between remotely sensed spatial patterns and some important characteristics of farming systems (Lambin, 1988; Gilruth and Hutchinson, 1990; Guyer and Lambin, 1993). Geoghegan et al. (1998) discuss two modelling approaches to discern information embedded within spatial imagery that is directly relevant to the core themes of the social sciences, and use it to inform the concepts and theories pertinent to those themes. In other words, this body of literature suggests that one can better understand how and why changes in land use are taking place from representations of the patterns of change, as observed for example by remote sensing and as described by spatial correlations between observed land-cover changes and proximate causes.

The objectives of this study are: (i) to develop a spatial, statistical model of the proximate causes of different processes of land-use change in the Mara Ecosystem, taking into account the spatial variability of the land-use change processes; and (ii) to identify how much understanding of the driving forces of these changes can be gained through such a spatial, statistical analysis.

2. Study area

The study area ($10,694\,\mathrm{km}^2$) is located in Narok District, in southwestern Kenya, between $34°45'$E and $36°00'$E at the Kenya–Tanzania border and adjacent to the Serengeti National Park in Tanzania. The rangelands surrounding the Masai Mara National Reserve can be divided into three range units based on bio-geographic and climatic differences (Fig. 1). The western unit consists mainly of grasslands and comprises the Masai Mara National Reserve. The Loita Plains stretch out in the northeastern part of the study area and are covered by dwarf shrub and whistling thorn (*Acacia drepanolobium*) grasslands. The eastern area, with the Siana Hills and Plains, supports *Croton dichogamus* bush and several other woody species interspersed with grasslands (Stelfox et al., 1986). North of Narok, the district capital, the Mau Uplands are the result of recent volcanic activity with highly porous and fertile volcanic ash soils (Said et al., 1997). The uplands are part of what is known in Kenya as the "high potential lands" and
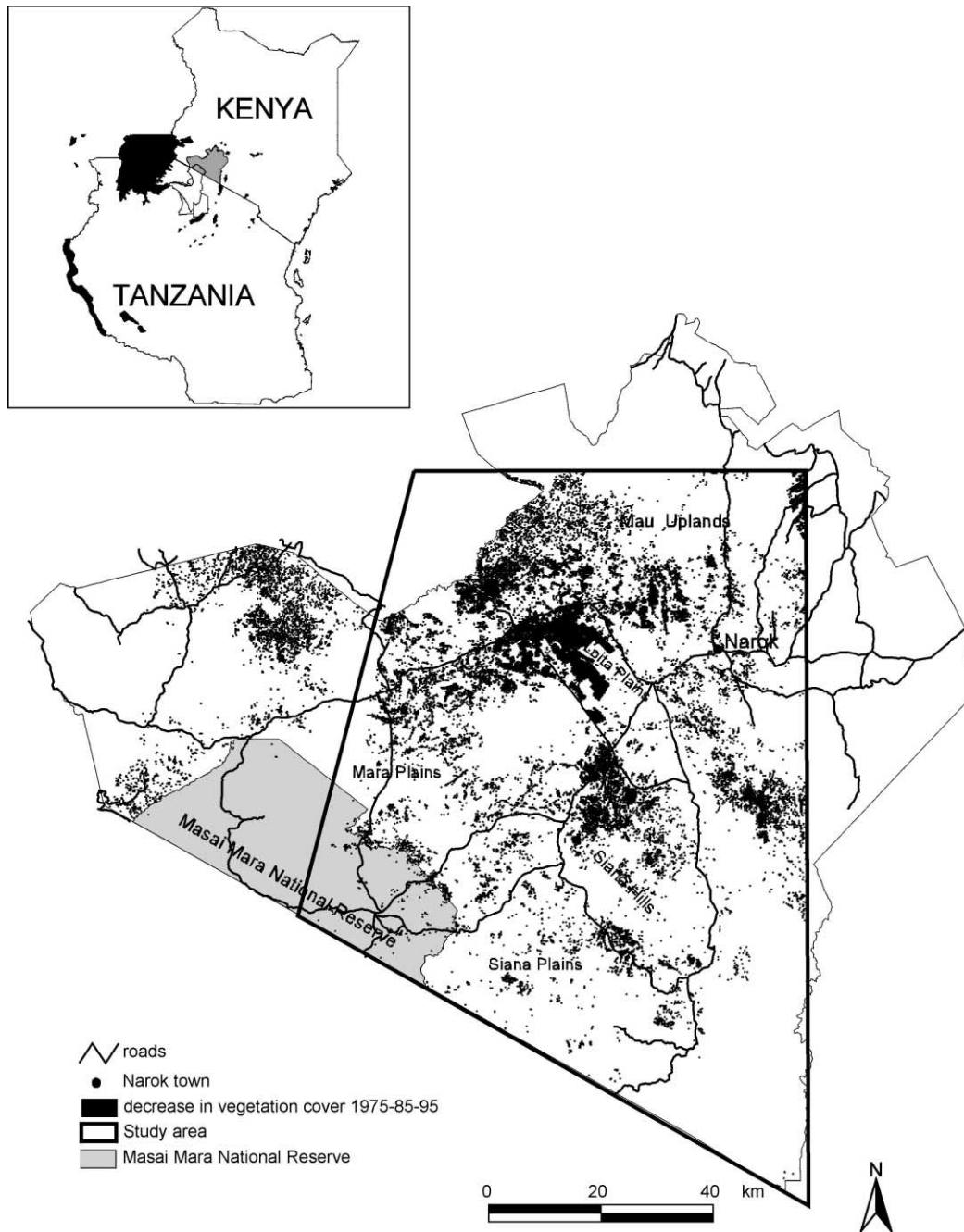
Fig. 1. Study area in Narok District, Kenya.

were extensively opened up for agriculture since the early 1960s (Singida, 1984).

The Masai Mara National Reserve (Fig. 1) was established in 1961. The 1523 km^2 reserve is a formal conservation estate, where land use is restricted to wildlife viewing tourism. The rangelands surrounding this area of fortress conservation were destined to act as a buffer zone between the Masai Mara National Reserve and the cultivated "high potential lands" in the northern part of Narok District. The rangelands surrounding the Mara Reserve contain year-round communities of resident wildlife, but migratory wildlife also spill out onto them during the dry season. As such, the grazing resources are important to the Maasai pastoralists, their livestock, and the wildlife alike. Since independence of the country, these rangelands were held as trust land by the Narok County Council. The 'Land Group Representatives and Land Adjudication Act' of 1968 enabled the land to be demarcated into group ranches, which were owned and under private title by a group of registered members and managed by an elected committee. Throughout the 1980s, more and more land within the group ranches was allocated to individuals and the tendency towards complete sub-division of the ranches gained more momentum (Singida, 1984; Galaty, 1992). Important changes in land use took place in the rangelands of the study area. Large-scale wheat farming in the area was initiated in the early 1970s, but knew its largest expansion from the eighties onward (Serneels et al., 2001). In Narok District, an area of 4875 ha in the northern part of the Loita eco-unit was cultivated in 1975 (Karime, 1990). Between 1975 and 1995, an additional 44,000 ha were converted to large-scale wheat farming in the Loita Plains (Serneels et al., 2001). Expansion continues at present. In total, 8.4% of the study area has been subject to a decrease in vegetation cover over the 20-year study period (Fig. 1), including changes due to mechanised agriculture, rangelands modification and smallholder agriculture.

Conversion to agriculture is the result of human decision-making processes (Fig. 2). Two types of agriculture are present: large-scale mechanised agriculture and smallholder farming. Expansion of the former is possibly driven by land suitability and economic factors, such as cereal and input prices, accessibility to the market, in this case to Narok, and transportation costs (Brush and Turner, 1987).
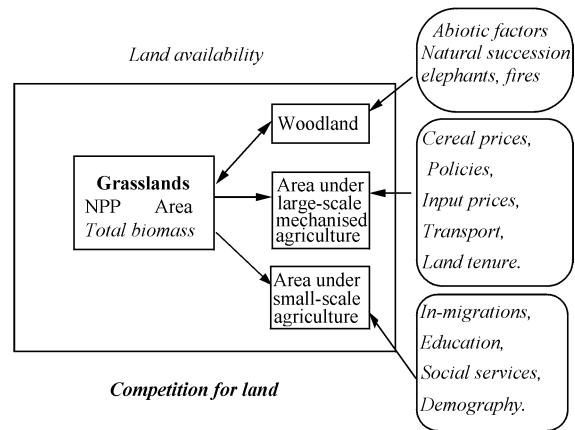

Fig. 2. Factors likely to influence land-use changes in the study area: competition between different land-use types and their potential driving forces.

Since individual landowners can decide how to use their land, Maasai landowners have had the option of maximising the development potential of their land by leasing it to farmers (Norton-Griffiths, 1996). Thus, changes in land tenure may have spurred the conversion of rangelands to agriculture.

The expansion of small-scale, smallholder agriculture is likely to be driven by different factors (Fig. 2), such as changes in demography, caused by in- or out-migration and population growth (Entwistle et al., 1998), and socio-economic factors such as education and social services. In the early years of the study period, the practise of small-scale agriculture has been linked with in-migrants from neighbouring districts (Lamprey, 1984). However, cultivation of small plots is becoming increasingly common, as is the development of permanent settlements (Lamprey and Waller, 1990). These changes might be partly linked with the introduction of individual and corporate titles to land instead of communal tenure.

The transition between woodland and grassland is largely driven by abiotic factors such as fires, heavy browsing by elephants or other large herbivores and natural successions (Sprugel, 1991; Ruess and Halter, 1990; Dublin et al., 1990; Turner et al., 1995), which themselves can be driven by human factors (e.g. human-caused burning, release of elephant populations, hunting pressure due to the ban on ivory trade).

3. Materials and methods

3.1. Data

A time series of three high resolution Landsat images was acquired for the years 1975 (Landsat-MSS), 1985 and 1995 (Landsat-TM). The study area was limited to the area common to the three Landsat images (Fig. 1). Change detection techniques were applied to image pairs and nine trajectories of change were mapped (Serneels et al., 2001). The change trajectories are based on changes in the normalised difference vegetation index (NDVI) of a given pixel. The NDVI was calculated from reflectances in the red and near infrared (NIR) parts of the electromagnetic spectrum as $(NIR-red)/(NIR+red)$. NDVI-image pairs of consecutive dates were subtracted. A pixel in the difference image is considered having been converted from one vegetation type to another if its difference in NDVI between the two dates is larger than a certain threshold value. The threshold values were chosen conservatively as to exclude the more subtle changes within vegetation types (Serneels et al., 2001). The resulting land-cover change map was validated in the field.

The roads network, rivers and villages were manually digitised from the 1:250,000 Survey of Kenya Topographic maps (1980), and updated with the recent satellite images. Human population distribution maps were derived from the Kenya Population Census of 1979 and 1989 (Republic of Kenya, 1979, 1989). The land adjudication map was digitised from the 1:250,000 Survey of Kenya map (1970). The soil map was digitised from the 1:1,000,000 exploratory soil map of Kenya, compiled by the Kenya Soil Survey in 1980 (Sombroek et al., 1982).

All data were brought together in a raster GIS and resampled to a common spatial resolution of 100 m. This resolution was largely determined by the nature of the dependent variables. While working at a coarser resolution might have better suited the quality of the independent variables, it would have altered the spatial patterns of fine-scale land-cover change that are to be modelled (Table 1).

3.2. Dependent variables

Most of the changes detected over the 20-year period are a decrease in vegetation cover. Three

Table 1
GIS database

Variables	Type	Unit
Dependent variables		
Mechanised agriculture, 1975–1985	Binary	0–1
Mechanised agriculture, 1985–1995	Binary	0–1
Smallholder impact, 1975–1985	Binary	0–1
Smallholder impact, 1985–1995	Binary	0–1
Rangeland modifications, 1985–1995	Binary	0–1
Independent variables		
Distance to Narok	Continuous	km
Distance to roads	Continuous	hm
Distance to villages	Continuous	hm
Distance to water	Continuous	hm
Elevation	Continuous	hm
Agro-climatic zone	Categorical	I–V
Soil suitability	Categorical	1–5
Land tenure — inner group ranch	Binary	0–1
Land tenure — outer group ranch	Binary	0–1
Population density, 1979	Continuous	in h./km^2
Population density, 1989	Continuous	in h./km^2
% difference in population density	Continuous	%

processes of land-cover change have been distinguished: (1) conversion to large-scale, mechanised wheat cultivation, (2) smallholder impact, i.e. a combination of clearing for subsistence agriculture and establishment of permanent settlements, and (3) rangeland modifications characterised by a loss of vegetation cover. Their spatial distribution was modelled separately for two periods: 1975–1985, and 1985–1995. The land-cover change map was segmented in zones with different dominant land-uses, based on the interpretation of the landscape patterns on the 1995 Landsat image and supported by field survey data.

3.3. Independent variables

Distance to roads: This variable was calculated as a series of buffers of 100 m (1 pixel) expanding from each arc of the road network. Most of the roads in the study area are gravel roads. Road quality largely depends on maintenance efforts and is highly variable in time. Each road segment was therefore treated as equally suitable for transport of goods and people.

Distance to towns: The distance to the nearest settlement was calculated as a series of buffers of 100 m, expanding from each centre. Only the officially

registered village centres were taken into account, as well as two of the game lodges that are located in the Masai Mara National Reserve. Although not real settlements, the lodges create job opportunities for the local people, as well as possibilities to trade locally made handicrafts. Equal weights were given to each village centre.

Distance to Narok: Narok is the district capital and the most important town in the region. It has the largest market in the area, hosts the government services and is the main link with Nairobi. As Narok is the economic heart of the region, the distance to the district capital was considered an important explanatory variable in the model. Distance to Narok was calculated as a series of buffers of 1 km, expanding from the district capital.

Distance to water: Accessibility to permanent water sources is highly valued by the largely pastoralist community. Livestock need regular access to

drinking water. Besides permanent rivers, permanent springs and artificial water holes with permanent water were also mapped. Temporary rivers and seasonal water points were not taken into account, as their seasonal importance is difficult to account for. Distance to water was calculated as a series of buffers of 100 m, expanding from each arc of the river or from each point in the case of springs and wells.

Land tenure: Private land tenure can be an important determinant in the decision-making process about land-use although, in our study region, land-use changes sometimes preceded the land subdivision process. Land tenure is represented as a set of binary layers, depicting sets of group ranches with similar land tenure system and similar development potential. The inner group ranches are the ones close to the Masai Mara National Reserve (Fig. 3a). The land is adjudicated in most cases, but still under communal management. The inner group ranches are important
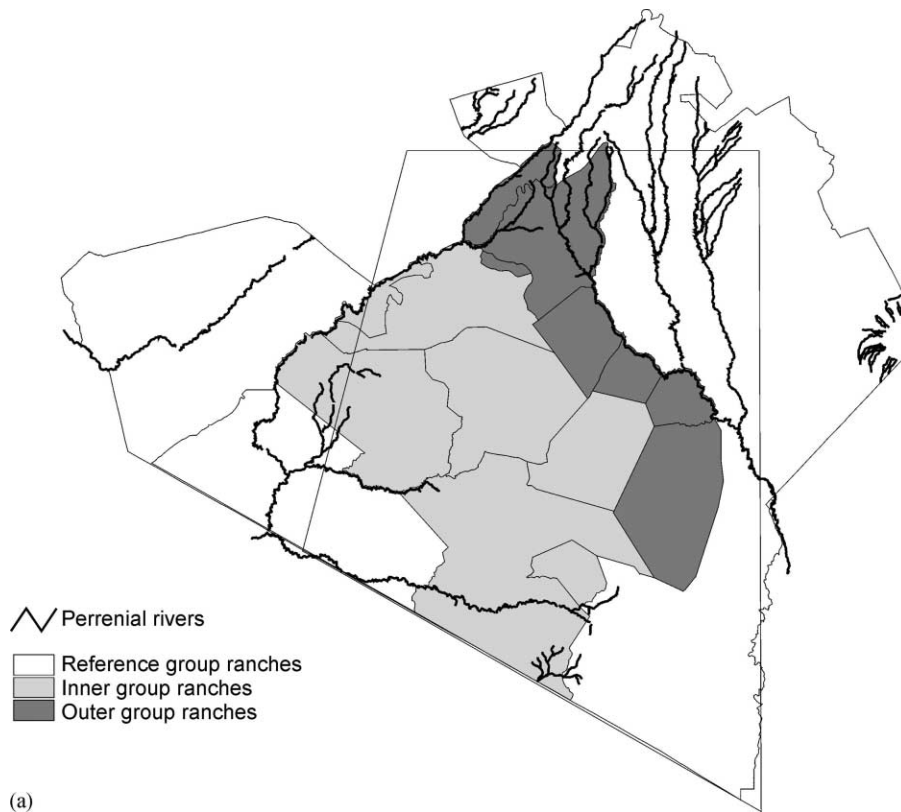

Perrenial rivers

Reference group ranches

Inner group ranches

Outer group ranches

(a)

Fig. 3. (a) Location of the inner and outer group ranches. (b) Agro-climatic zones in Narok District.

roads
settlement
Agro-climatic zones
I (humid)
II (sub-humid)
III (semi-humid)
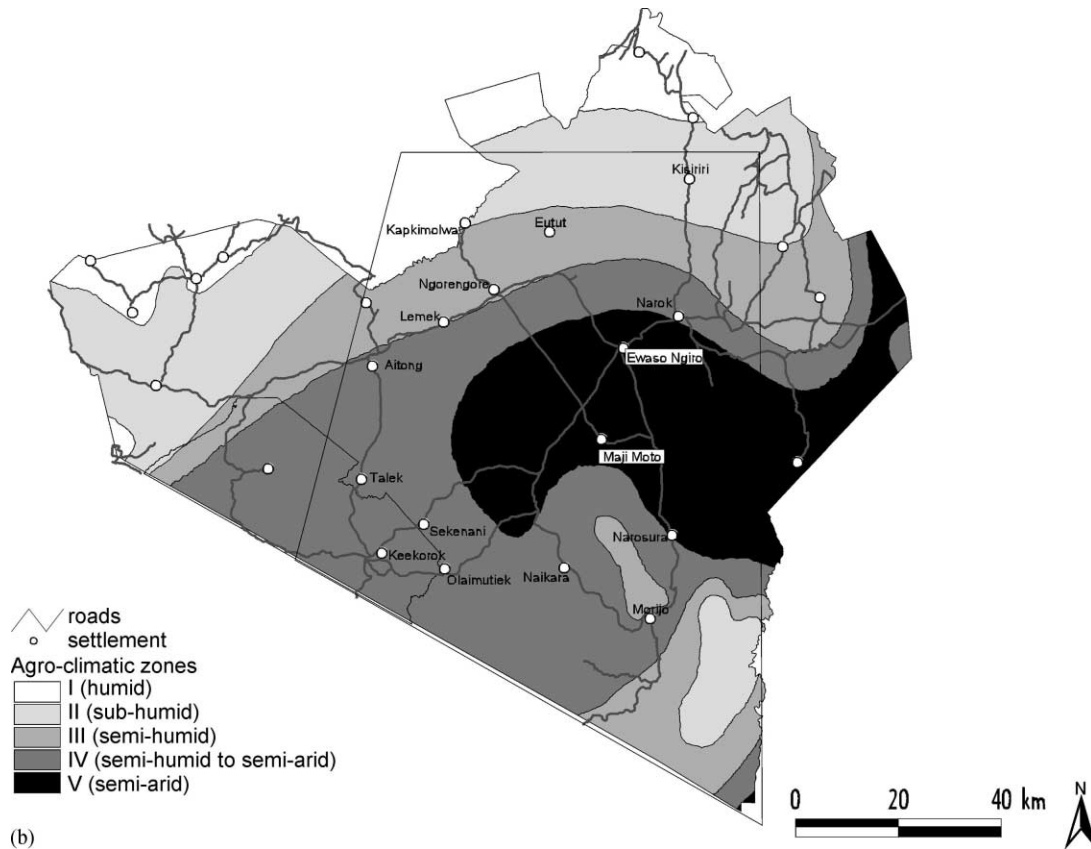IV (semi-humid to semi-arid)
V (semi-arid)

(b)

Fig. 3 (*Continued*).

as wildlife dispersal areas and have a high development potential for tourist-related activities. The outer group ranches form a ring around the inner group ranches. They are further away from the Masai Mara National Reserve and thus have a lower development potential for tourist-related activities. Most of the outer group ranches have been subdivided in private land titles. The remaining group ranches are used as the reference group.

Altitude: A Digital Elevation Model was constructed from the contour lines present in the Arc/Info Digital Chart of the World (ESRI, 1993). The contour lines were digitised at the 1:1,000,000 scale, at intervals of 1000 ft. The resulting elevation map was reclassified in units of 100 m intervals.

Agro-climatic zone: The agro-climatic zoning system (Sombroek et al., 1982) is based on a combination of two sets of environmental variables: (i) the moisture availability zones, which are based on the ratio of the measured average annual rainfall and the calculated average annual potential evaporation, and (ii) the temperature zones, which are based on the average annual temperature and the altitude. The map of agro-climatic zones represents the potential agricultural suitability of a region, with consideration of moisture and temperature constraints, but not taking into account possible limitations arising from the soil quality. The map of agro-climatic zones for Narok District represents five classes, with I: humid, II: sub-humid, III: semi-humid, IV: semi-humid to semi-arid, V: semi-arid (Fig. 3b).

Suitability for agriculture: The exploratory soil map for Narok District (Sombroek et al., 1982) was remapped into five suitability classes. These classes (from "1": least suitable, to "5": most suitable) were based on the intrinsic soil characteristics of the soil types, such as water retention capacity, sandiness,

organic matter content, etc. The potential suitability for agriculture of a soil type under appropriate management practises was not taken into account.

Population density: Three population layers were used in the models: population density at sub-location level for 1979 and 1989, and difference in population density between those dates. As the enumeration units changed between censuses, sub-locations were aggregated to a location level, so that the aggregated enumeration areas for the 1979 and 1989 censuses had a maximum overlap. Percentages of change in population density from 1979 to 1989 were calculated for these aggregated units.

3.4. Multiple logistic regression models

The technique used in this study is multiple logistic regression (MLR). It is designed to estimate the parameters of a multivariate explanatory model in situations where the dependent variable is dichotomous, and the independent variables are continuous or categorical. The MLR technique yields coefficients for each independent variable based on a sample of data. These coefficients are interpreted as weights in an algorithm that generates a map depicting the probability of a specific category of land-use change for all sampling units. MLR has already been successfully used in wildlife habitat studies (Pereira and Itami, 1991; Narumalani et al., 1997; Bian and West, 1997), the prediction of forest fires (Vega Garcia et al., 1995), and deforestation analyses (Ludeke et al., 1990; Chomitz and Gray, 1996; Mertens and Lambin, 2000).

MLR identifies the role and intensity of explanatory variables X_n in the prediction of the probability of one state of the dependent variable, which is defined as a categorical variable Y. Suppose X is a vector of explanatory variables and p is the response probability to be modelled with, in the case of a dichotomous dependent variable, $p = \Pr(Y = 1|X)$, with $Y = 0$ meaning the absence of agriculture and $Y = 1$ meaning the presence of agriculture. The linear logistic model has the form

$$\text{logit}(p) = \log \left[\frac{p}{1-p} \right]$$
$$= \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \quad (1)$$

where α is the intercept and β_n are slope parameters.

The probability values can thus be quantitatively expressed in terms of explanatory variables by

$$p = \frac{\exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n)}{1 + \exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n)} \quad (2)$$

Odds ratios can be used to facilitate model interpretation (Stokes et al., 1995; Menard, 1995). The odds ratio ($\psi = e^{\beta i}$) is a measure of association which approximates how much more likely (or unlikely) it is for the outcome to be present for a set of values of independent variables (Hosmer and Lemeshow, 1989). The odds ratio can be interpreted as the change in the odds for an increase of one unit in the corresponding risk factor. To calculate the odds ratio for a change in the risk factor with more than one unit, e.g. from a to b units, the odds ratio estimate is raised to the power $c = b - a$ (SAS Institute, 1997). The probability, the odds and the logit are three different ways of expressing the same thing (Menard, 1995). The estimated odds values are computed as the exponential of the parameter estimate values (Hosmer and Lemeshow, 1989; Agresti, 1990):

$$\text{odds}(p) = \exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n) \quad (3)$$

In this study, MLR was performed using the LOGISTIC function in the SAS/STAT software. The predictive ability of a logistic regression model is evaluated from the table of maximum likelihood estimates (MLE), which contains the MLE of the parameters, the estimated standard errors of the parameter estimates, the Wald χ^2 statistics and the significance probabilities for the parameter estimates. Positive values of the parameter estimate indicate that larger values of the explanatory variable will increase the likelihood of the occurrence of the event. Likewise, negative values of the parameter estimate indicate that larger values of the explanatory variable will decrease the likelihood of the occurrence of the event. The χ^2 statistic indicates the relative weight of each explanatory variable in the model and allows us to assess the role of each variable in the prediction of an event.

In the case of logistic models, the goodness-of-fit measure is defined as the ratio of maximised log likelihood. This pseudo-R^2 or ρ^2 is defined as (Wrigley, 1985)

$$\rho^2 = 1 - \frac{\log[\beta]}{\log[C]} \quad (4)$$

that is 1 minus the ratio of the maximised log likelihood values of the fitted log[β] and constant-only-term log[C] models. Although ρ^2 ranges in the value from 0 to 1, its value tends to be considerably lower than the value of the coefficient of determination R^2 of conventional regression analysis. It should not be judged by the standards of what is normally considered a "good fit" in conventional regression analysis (Wrigley, 1985). Values between 0.2 and 0.4 should be taken to represent a very good fit of the model (Domencich and McFadden, 1975).

To fit the model, several options are available, such as stepwise regression, "best subset" models or predefined conceptual models. In this study, the latter approach was chosen. In the statistical model, we introduced all explanatory variables suggested by the conceptual model (Fig. 2) and, based on the full model information, analysed which variables contribute significantly to the explanation of land-use changes.

3.5. Sampling procedure

Prior to performing the MLR, a stratified random sampling procedure was used to select N observation points distributed throughout the study area. Random sampling of observations was used due to the presence of spatial auto-correlation in the data. For every sample observation, the values of the dependent and the set of independent variables were recorded. For each model, a random sample of 10,000 observations was selected (0.58% of the study area), with an equal number of 0 and 1 observations of the dependent variable. Unequal sampling rates do not affect the estimation of the coefficients of the explanatory variables in logit models (Maddala, 1988), but only affect the constant term. When using the model to run simulations, the constant term is corrected by adding ($\ln p_1 - \ln p_2$), where p_1 and p_2 are the proportions of observations chosen from the two groups for which the dependent variable is 1 and 0, respectively (Maddala, 1988).

The independent variables in the samples were tested for multicollinearity. The linearity of the bivariate relationship between each of the independent variables and the dependent variable was tested (Hosmer and Lemeshow, 1989). Continuous variables that did not exhibit linear behaviour were transformed and logarithmic forms or squared terms were used in the MLR models. The ordinal categorical variables were transformed into sets of binary variables. Depending on the model, these sets of binary variables needed regrouping. This is mentioned in the results section for each of the models.

4. Results

4.1. Preliminary statistical analysis

Low levels of collinearity between the independent variables were found. The coefficients of determination (R^2) of the multivariate relationships between one of the independent variables against all the others range from 0.12 to 0.75, which is below the critical value of 0.80 (Menard, 1995). Thus, all independent variables were used in the MLR analysis.

4.2. Spatial model for mechanised agriculture

In the models for mechanised agriculture, agro-climatic zones were represented by two dummy variables representing zones IV and V, respectively. Agro-climatic zones I–III were taken as reference. Soil suitability was represented as a binary variable, with soils less suitable for agriculture (values 1–3) coded 0 and soils suitable for agriculture (values 4 and 5) coded 1.

The models for the first and second periods both have a very good overall explanatory power, with ρ^2 of 0.63 and 0.51, respectively (Table 2a and b). The two most important explanatory variables in both models are the agro-climatic zones and distance to Narok. In both models, the probability of finding mechanised agriculture increases for more suitable agro-climatic zones. The odds of finding agriculture in zone IV compared to the reference zone, are 0.127, i.e., agriculture is 7.9 times more likely to be found in the reference zone than in agro-climatic zone IV. In the first model, agriculture has a very low probability of occurring in the semi-arid agro-climatic zone V. However, the probability of finding mechanised agriculture in zones IV and V increased in the second period. The likelihood of observing mechanised agriculture increases as the distance to Narok decreases in both models. For each additional 10 km away from Narok, it is five times less likely to find mechanised

Table 2

Expansion of mechanised agriculture: (a) between 1975 and 1985, $\rho^2 = 0.63$; (b) between 1985 and 1995, $\rho^2 = 0.51$

Variable	Parameter estimate	Standard error	Wald χ^2	Pr $> \chi^2$	Odds ratio
(a) Between 1975 and 1985, $\rho^2 = 0.63$					
Intercept	20.525	0.938	478.9	0.0001	–
Agro-climatic zone IV	−2.061	0.110	351.8	0.0001	0.127
Agro-climatic zone V	−8.029	0.214	1411.5	0.0001	0.0003
Distance to Narok	−0.161	0.005	933.7	0.0001	0.851
Elevation	−0.714	0.043	276.0	0.0001	0.49
Population density, 1979	−0.0173	0.001	271.1	0.0001	0.983
Soil suitability	−1.523	0.120	160.4	0.0001	0.218
Inner group ranches	1.565	0.250	39.3	0.0001	4.783
Outer group ranches	2.234	0.181	151.7	0.0001	9.339
Distance to village	−0.0103	0.00137	56.6	0.0001	0.99
Distance to water (log)	0.273	0.0435	39.2	0.0001	1.313
Distance to road	0.0255	0.00410	38.7	0.0001	1.026
(Distance to road)^2	−0.00018	0.000038	22.6	0.0001	1
(b) Between 1985 and 1995, $\rho^2 = 0.51$					
Intercept	13.240	0.707	351.1	0.0001	–
Agro-climatic zone IV	−1.216	0.105	135.1	0.0001	0.297
Agro-climatic zone V	−5.073	0.161	987.7	0.0001	0.006
Distance to Narok	−0.122	0.00447	745.3	0.0001	0.885
Elevation	−0.712	0.0342	433.8	0.0001	0.491
Distance to water (log)	0.733	0.0417	309.2	0.0001	2.082
Inner group ranches	2.925	0.186	248.0	0.0001	18.640
Outer group ranches	2.401	0.144	277.5	0.0001	11.031
% difference in population density	0.00707	0.000502	198.1	0.0001	1.007
Population density, 1989	0.0121	0.000863	196.8	0.0001	1.012
Distance to village	0.00835	0.000877	90.7	0.0001	1.008
Distance to road	0.00904	0.00349	6.7	0.0096	1.009
(Distance to road)^2	−0.00019	0.000036	28.5	0.0001	1
Soil suitability	−0.244	0.0958	6.5	0.0108	0.783

agriculture (10 km = 10 units thus the odds ratio is $e^{-0.161 \times 10} = 0.2$; Table 2a).

In both models, the probability of observing mechanised agriculture decreases with increasing elevation. Land tenure (inner and outer group ranch variables) has good explanatory power in both models and reflects the expansion of mechanised agriculture in the outer group ranches in the first period and in the inner group ranches in the second period. The probability of finding mechanised agriculture in both models increases with increasing distance to permanent water sources. The estimated odds to find agriculture at 1 km from a water source are 1.3 and 2.1 times higher than at 100 m from a water source for the first and second periods, respectively. Whereas the probability of mechanised agriculture increased with proximity to villages in the first period, the relationship is reversed

in the second period. The large-scale mechanised agriculture is expanding away from the settled areas. Agriculture is likely to develop on less suitable soils in the first period and soil suitability has little explanatory power in the second period. Thus farming areas are not selected for their soil properties. In the first period, mechanised agriculture is explained by lower population density in 1979, whereas it is explained by a higher population density in 1989 and increased population density since 1979 in the second period.

4.3. Spatial model for smallholder impact

In the models for smallholder impact, agro-climatic zones were introduced as three binary variables for agro-climatic zones III–V. Zones I and II were taken as reference. Soil suitability was also used as a set

Table 3
Smallholder impact between 1975 and 1985, $\rho^2 = 0.20$

Variable	Parameter estimate	Standard error	Wald χ^2	Pr > χ^2	Odds ratio
Intercept	6.759	0.451	224.5	0.0001	–
Agro-climatic zone V	−3.446	0.143	579.3	0.0001	0.032
Agro-climatic zone IV	−2.399	0.116	427.7	0.0001	0.091
Agro-climatic zone III	−0.718	0.099	52.5	0.0001	0.488
Distance to water (log)	−0.274	0.0253	117.3	0.0001	0.76
Inner group ranches	0.413	0.0733	31.7	0.0001	1.511
Outer group ranches	−0.765	0.0821	86.8	0.0001	0.466
Elevation	−0.159	0.0176	81.8	0.0001	0.853
Distance to village (log)	−0.252	0.0446	31.9	0.0001	0.777
Distance to road (log)	−0.107	0.0228	22.1	0.0001	0.898
Distance to Narok	0.00924	0.00197	22.0	0.0001	1.009
Soil suitability-3	0.397	0.0754	27.7	0.0001	1.487
Soil suitability-4	−1.170	0.364	10.3	0.0013	0.31
Soil suitability-5	−0.282	0.0897	9.9	0.0017	0.754
Population density, 1979	−0.00018	0.00064	0.1	0.7738	1

of binary variables, with categories 1 and 2 (very low suitability) as reference and binary layers for categories 3–5. The models for smallholder impact have a lower explanatory power than the ones for mechanised agriculture (Table 3). The first model has a reasonably good fit, with $\rho^2 = 0.20$, but the second model is not significant ($\rho^2 = 0.12$).

In the first period, smallholder impact is primarily explained by agro-climatic potential, with more negative parameter estimates for the dryer zones.

Proximity to water is also an important explanatory variable. At 100 m from a permanent water source, it is 1.3 times more likely to find smallholder impact than at 1 km from that source. Smallholder impact was more likely to occur in the inner group ranches and less likely to occur in the outer group ranches than elsewhere in the study area. The probability for smallholder impact increased with decreasing altitude. Proximity to villages and roads are also explaining smallholder impact in the first model.

Table 4
Rangelands modification between 1985 and 1995, $\rho^2 = 0.60$

Variable	Parameter estimate	Standard error	Wald χ^2	Pr > χ^2	Odds ratio
Intercept	−18.089	0.957	357.4	0.0001	–
(Distance to Narok)^2	−0.00652	0.000283	531.9	0.0001	0.993
Distance to Narok	0.421	0.0225	350.4	0.0001	1.524
Population density, 1989	−0.195	0.0114	295.2	0.0001	0.823
Distance to road	0.0196	0.00139	199.3	0.0001	1.02
% difference in population density	−0.0139	0.00109	162.9	0.0001	0.986
Agro-climatic zone V	3.562	0.298	142.5	0.0001	35.222
Agro-climatic zone IV	2.210	0.282	61.5	0.0001	9.116
Elevation	0.362	0.035	109.5	0.0001	1.436
Inner group ranches	2.268	0.233	94.4	0.0001	9.655
Outer group ranches	0.767	0.206	13.9	0.0002	2.154
Soil suitability-3	1.815	0.209	75.4	0.0001	6.142
Soil suitability-4	1.021	0.217	22.2	0.0001	2.776
Soil suitability-5	0.154	0.114	1.8	0.1783	1.166
Distance to water (log)	0.524	0.0612	73.4	0.0001	1.689
Distance to village	0.00348	0.000833	17.4	0.0001	1.003

4.4. Spatial model for rangeland modifications

In the rangeland modifications model, agro-climatic zones were represented by two binary variables for zones IV and V. Zones I–III were taken as reference. Soil suitability was also used as a set of binary variables, with categories 1 and 2 (very low suitability) as reference and binary layers for categories 3–5. Between 1975 and 1985, rangeland modifications were not observed in the study area. The overall explanatory power of the model for rangeland modi-

fications from 1985 to 1995 is very high, with a ρ^2 of 0.60 (Table 4). The presence of the squared term for distance to Narok indicates that there is an optimal distance interval for rangeland modifications. It also points to the presence of spatial auto-correlation, with adjacent areas more likely to undergo the same type of land-cover changes. Rangeland modifications are likely to occur in sparsely populated areas with low rates of population increase. Finally, the probability for rangeland modifications increases in the dryer agro-climatic zones and further away from roads.
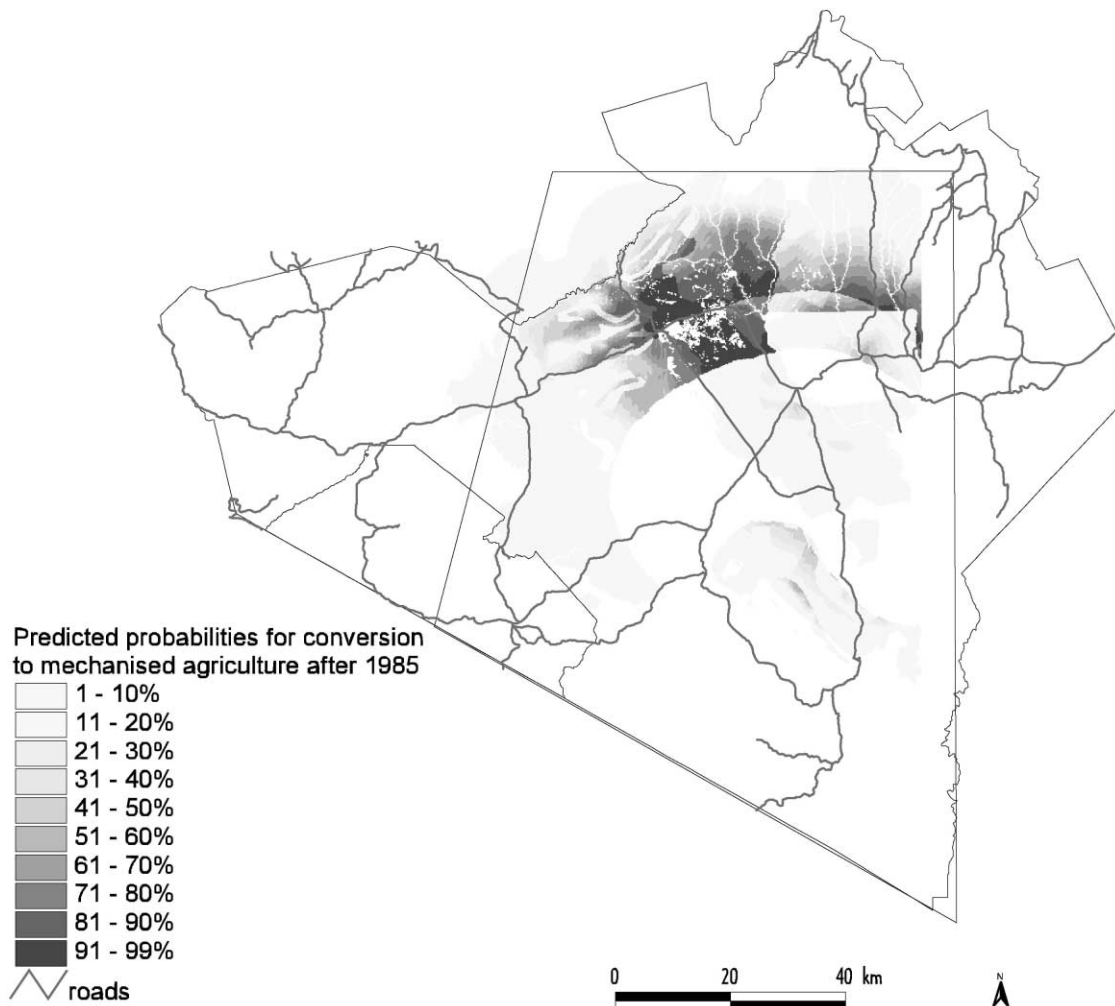

Fig. 4. Probability map for conversion to mechanised agriculture after 1985, based on the model for 1975–1985. Locations already cultivated in 1985 are masked.

Each additional kilometre away from roads increases the odds of finding rangeland modifications by a factor of 1.21.

4.5. Validation

In the case of mechanised agriculture, the same explanatory variables appear for both periods, although they do not have the same weights. We tested how accurately the areas that were likely to be converted to mechanised agriculture in the second period could be predicted, using the parameter estimates of the model for the first period in Eq. (2). A probability surface associated with mechanised agriculture was thus generated. The map represents the areas that are at risk of being converted to mechanised agriculture after 1985, assuming that similar causal processes as observed between 1975 and 1985 would be maintained (Fig. 4). The predicted probabilities of conversion to mechanised agriculture for areas not under agriculture in 1985 were compared with the map of actual conversions between 1985 and 1995. Overall, less than 10%

of the pixels with a predicted probability below 50% were converted to agriculture (Fig. 5). Up to 50% of all the pixels with a predicted probability for conversion to agriculture higher than 90% were converted to mechanised agriculture after 1985. However, a second peak of observed conversions was found for predicted probabilities of 70–85%, with about 15–20% of these pixels actually being converted to agriculture between 1985 and 1995. These pixels are located in the inner group ranches or in areas with soils highly suitable for agriculture. Additionally, some expansion of agriculture into the semi-arid agro-climatic zone V took place after 1985. The parameter estimates for these three variables in the 1975–1985 model are such that expansion of agriculture in either area after 1985 was predicted with a lower probability. Actually, before 1985, little expansion of agriculture took place in the inner group ranches, in areas with highly suitable soils for agriculture or in the semi-arid zones. The ability of the first-period model to predict land-use changes in the second period is thus limited.
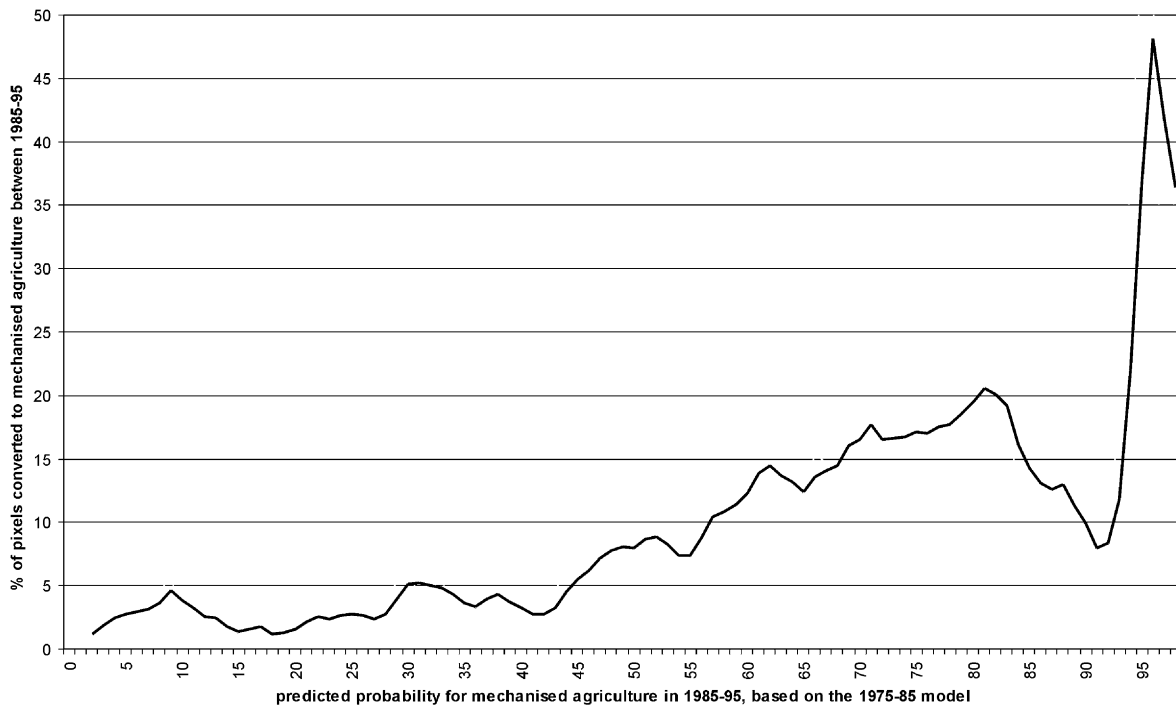

Fig. 5. Percentage of pixels with a given predicted probability that have actually been converted to agriculture between 1985 and 1995.

5. Discussion

Land-cover changes in Narok District are the result of a variety of processes of land-use change. For mechanised agriculture, agro-climatic potential and accessibility to the market are the most important explanatory variables. Low altitude plains are preferred, which are easily accessible with heavy machinery. Accessibility is more important than soil quality, as the latter can be improved with fertilisers. Also, the immediate surroundings of water sources are not cultivated. Hence, these findings indicate that pastoralists preferentially lease those lands that are further away from water sources, or that are less fertile, and thus have a lower rent value for them. While conversion in the first period mainly took place in the outer group ranches, which were mostly under private land title, expansion of agriculture into the inner group ranches and into areas with more fertile soils happened mainly in the second period. The inner group ranches were communally owned, but outsiders could negotiate leases with powerful group ranch leaders (Thompson, in preparation). The increasing population density associated with mechanised farming activities in the second period raises several hypotheses. More densely populated areas might become less suitable to practise extensive cattle raising and thus people look for alternative incomes. Alternatively, the revenues gained from mechanised agriculture might attract people from other regions, thus causing a stronger population increase in the area.

The models of smallholder impact represent two distinct processes: the development of permanent settlements and expansion of smallholder agriculture. In the first period, the former process is much more important and the spatial model represents landscape attributes that are attractive for settlement. Pastoralists value the vicinity to water for their livestock. They also value the proximity to the Masai Mara National Reserve (inner group ranches), because of the availability of permanent water (i.e. the Mara and Talek Rivers), the possibility of supplementing their income by tourist-related activities and the presence of good grazing grounds in the National Reserve. Although grazing by cattle in the National Reserve is generally not allowed, the Maasai pastoralists are granted temporary access to the park in times of drought. Hence, some permanent settlements developed at the park

borders, close to the Talek River, or close to the park entrance gates (Sekenani and Olaimutiek), along the access roads to the park.

In the second period, subsistence agriculture became more important and the model parameters represent, on the one hand, expansion of smallholder agriculture in the more fertile areas with higher elevation and, on the other hand, the continuing process of settlement development in the inner group ranches. The concomitant modelling of these two simultaneous processes results in a weak explanatory power of the overall model. Further subdivision of the region into smaller homogeneous spatial entities, separating these two processes of land-use change, might raise the predictive power of the individual models. However, it is not compatible with the idea of generality behind modelling activities and it would make extrapolation of the findings to the entire region even more difficult.

Rangeland modifications (loss of vegetation cover) take place in the marginal lands, with low rainfall and with a high sensitivity to inter-annual rainfall variability. There is no evidence of increased wildebeest or livestock densities in these rangelands in the wet season (Serneels and Lambin, in press). It is thus unlikely that increased grazing pressure on the rangelands in the wet season has led to these rangeland modifications. However, further research is needed to investigate the impacts of a possible increased grazing pressure during the dry season, increased sheep and goats densities, and the fencing of some private lands in this area.

The above results were obtained at a spatial resolution of 100 m. The results are valid for the spatial resolution and extent applied in the model. The choice of the scale, spatial extent and resolution for this study critically affected the patterns of land-use change that were observed (Gibson et al., 2000). Extending the spatial extent of the study area might have included land-use change processes that are currently not addressed, while changing the spatial resolution of the analysis would have altered the spatial patterns of land-use change. While patterns of land-cover change caused by the expansion of large-scale mechanised farming remain similar at coarser resolutions of, e.g. 1 km, the patterns of land-cover change linked to smallholder agriculture would no longer be detected at such spatial resolution. Other researchers have

demonstrated that explanatory variables change as the scale of analysis changes (e.g. Walsh et al., 1999; Kok et al., 2001).

The variability in time of land-use change processes affects our ability to use regression models for wide ranging extrapolations. A regression model that fits well in a region of the variable space corresponding to the original data can perform poorly outside that region. This is demonstrated in the validation exercise of the spatial model for mechanised agriculture, where the model for 1975–1985 performed moderately well in predicting the expansion of mechanised agriculture after 1985. Although the same explanatory variables re-occurred in the second model, some variables had evolved to more extreme values. This temporal heterogeneity is related, in this case, to the spatial diffusion of agriculture into different physiographic units. It could also be caused by changes in policies, macro-economic transformations or abrupt climatic events.

Can we uncover driving forces of land-use change through a statistical analysis of its proximate causes? Yes, the descriptive spatial models developed in this study suggest some important underlying causes driving the land-use changes that can be related to some well-established theoretical frameworks. For example, the explanatory variables of the spatial model of mechanised agriculture clearly suggest a von Thünen-like model, where conversion to agriculture is controlled by the distance to the market, as a proxy for transportation costs, and agro-climatic potential (von Thünen, 1966). Expansion of smallholder agriculture and settlements are also controlled by land rent. However, in this case, the rent is determined by other variables: proximity to permanent water, land suitability defined in terms of soil aptitude, location near a tourism market (i.e. the gates of the Masai Mara National Reserve) and vicinity to villages to gain access to social services (e.g. health clinics, schools, local market). This difference in perception of land rent reflects the widely different social and economic activities and objectives of smallholders versus the large entrepreneurs involved in mechanised farming. Finally, rangeland modifications are explained by biophysical factors determining the marginality of some rangelands and their sensitivity to climate fluctuations. This third process of land-use change suggests that a disequilibrium model of

ecological changes (Ellis and Swift, 1988) might be appropriate to account for such changes.

More fundamental driving forces such as land-use policies and land tenure are indirectly reflected in the model through the group ranch variable. However, assessing the spatial implications of government policies and land adjudication on land use in the area would require a detailed analysis of these policies, their timing, their actual implementation and their geographic impact. This would have to be supplemented by household surveys to understand how these exogenous policies have influenced decision-making processes by pastoralists.

6. Conclusion

Land-cover and land-use changes in Narok District are dynamic and diverse. The spatial distribution of the different types of land-cover or land-use changes can be modelled based on a set of cultural and natural landscape attributes that represent proximate causes of change. These models allow us to better understand which interactions between spatial determinants prevail to explain the spatial distribution of different land-cover change processes. However, it is difficult to distinguish between correlation and causality. For example, the mechanised agriculture model in the second period has population density (in 1989) and increase in population density as important explanatory variables. It is difficult to determine the direction of the causal relationship between population increase on the one hand, and expansion of agriculture on the other.

The spatial models developed here allow us to predict *where* land-cover conversions are most likely to take place in the near future. The models do not address the question "*when* are land-cover conversions likely to take place?" Even though they suggest some driving forces to explain these changes. To transform the map of predicted probabilities of land-cover change in actual land-cover change projections, a dynamic model that considers the driving forces of change needs to be developed. These driving forces (e.g. new government policies or changes in the prices of agricultural inputs or products) are often remote in space and/or time. However, the impact of changes in driving forces that are well represented by proximate causes, such as transport costs to the market, can be

evaluated with the spatial model. For example, the possible impact of the construction of a new road can be demonstrated in a map with associated predicted probabilities of land-cover change. The spatial models developed in this study can thus easily be used as simulation tools to support environmental impact assessments, decision-making and policy formulation.

Acknowledgements

References

Agresti, A., 1990. Categorical Data Analysis. Wiley, New York.

Bian, L., West, E., 1997. GIS modelling of elk calving habitat in a prairie environment with statistics. Photogramm. Eng. Remote Sens. 63, 161–167.

Brush, S.B., Turner, B.L., 1987. The nature of farming systems and views of their change. In: Turner, B.L., Brush, S.B. (Eds.), Comparative Farming Systems. Guilford Press, New York, pp. 11–48.

Chomitz, K.M., Gray, D.A., 1996. Roads, land use, and deforestation: a spatial model applied to Belize. World Bank Econ. Rev. 10, 487–512.

Domencich, T.A., McFadden, D., 1975. Urban Travel Demand: Behavioural Analysis. North-Holland, Amsterdam.

Dublin, H.T., Sinclair, A.R.E., McGlade, J., 1990. Elephants and fire as causes of multiple stable states in the Serengeti–Mara woodlands. J. Anim. Ecol. 59, 1147–1164.

Ellis, J.E., Swift, D.M., 1988. Stability of African pastoral systems: alternate paradigms and implications for development. J. Range Manage. 41, 450–459.

Entwistle, B., Walsh, S.J., Rindfuss, R.R., Chamratrithirong, A., 1998. Land-use/land-cover and population dynamics, Nang Rong, Thailand. In: Liverman, D., Moran, E.F., Rindfuss, R.R.,

Stern, P.C. (Eds.), People and Pixels: Linking Remote Sensing and Social Science. National Academy Press, Washington, DC, pp. 121–144.

ESRI, 1993. Digital chart of the world. Environmental Systems Research Institute, Inc., Redlands, CA.

Galaty, J.G., 1992. "The land is yours": social and economic factors in the privatisation, sub-division and sale of Maasai ranches. Nomadic Peoples 30, 26–40.

Geoghegan, J., Pritchard, L., Ogneva-Himmelberger, Y., Chowdhury, R.R., Sanderson, S., Turner II, B.L., 1998. "Socializing the pixel" and "pixelizing the social" in land-use and land-cover change. In: Liverman, D., Moran, E.F., Rindfuss, R.R., Stern, P.C. (Eds.), People and Pixels: Linking Remote Sensing and Social Science. National Academy Press, Washington, DC, pp. 51–69.

Gibson, C.C., Ostrom, E., Ahn, T.K., 2000. The concept of scale and the human dimensions of global change: a survey. Ecol. Econ. 32, 217–239.

Gilruth, P.T., Hutchinson, C.F., 1990. Assessing deforestation in the Guinea Highlands of West Africa using remote sensing. Photogramm. Eng. Remote Sens. 56, 1375–1382.

Guyer, J., Lambin, E.F., 1993. Land use in an urban hinterland: ethnography and remote sensing in the study of African intensification. Am. Anthropol. 95, 839–859.

Hosmer, D.W., Lemeshow, S., 1989. Applied Regression Analysis. Wiley, New York.

Karime, O.N.I., 1990. The changing land use patterns in the Ngorengore Loita Plains. Department of Resource Surveys and Remote Sensing, Ministry of Planning and National Development, Kenya.

Kok, K., Farrow, A., Veldkamp, A., Verburg, P.H., 2001. A method and application of multi-scale validation in spatial land use models. Agric. Ecosys. Environ. 85, 223–238.

Lambin, E.F., 1988. L'apport de la télédétection dans l'étude des systèmes agraires d'Afrique: l'exemple du Burkina Faso. Africa 58, 337–352.

Lambin, E.F., 1994. Modelling deforestation processes: a review. TREES Publications Series B: Research Report No. 1. European Commission, EUR 15744 EN.

Lamprey, R.H., 1984. Maasai impact on Kenya Savannah vegetation: a remote sensing approach. Ph.D. Thesis. University of Aston, Birmingham, UK.

Lamprey, R., Waller, R., 1990. The Loita–Mara region in historical times: patterns of subsistence, settlement and ecological change. In: Robertshaw, P. (Ed.), Early Pastoralists of Southwestern Kenya. Memoirs of the British Institute in Eastern Africa, Nairobi.

Liverman, D., Moran, E.F., Rindfuss, R.R., Stern, P.C. (Eds.), 1998. People and Pixels: Linking Remote Sensing and Social Science. National Academy Press, Washington, DC.

Ludeke, A.K., Maggio, R.C., Reid, L.M., 1990. An analysis of anthropogenic deforestation using logistic regression and GIS. J. Environ. Manage. 31, 247–259.

Maddala, G.S., 1988. Introduction to Econometrics. Macmillan, New York.

Menard, S., 1995. Applied Logistic Regression Analysis. Quantitative Applications in the Social Sciences, No. 106. Sage, London.

Mertens, B., Lambin, E.F., 1997. Spatial modelling of deforestation in southern Cameroon: spatial disaggregation of diverse deforestation processes. Appl. Geogr. 17, 143–162.

Mertens, B., Lambin, E.F., 2000. A spatial model of land-cover change trajectories in a frontier region in southern Cameroon. Ann. Assoc. Am. Geogr. 90, 467–494.

Narumalani, S., Jensen, J.R., Althausen, J.D., Burkhalter, S., Mackey Jr., H.E., 1997. Aquatic macrophyte modelling using GIS and multiple logistic regression. Photogramm. Eng. Remote Sens. 63, 41–49.

Norton-Griffiths, M., 1996. Property rights and the marginal wildebeest: an economic analysis of wildlife conservation options in Kenya. Biodiversity Conserv. 5, 1557–1577.

Pereira, J.M.C., Itami, R.M., 1991. GIS-based habitat modelling using logistic multiple regression: a study of the Mt. Graham red squirrel. Photogramm. Eng. Remote Sens. 57, 1475–1486.

Pontius Jr, R.G., Schneider, L.C., 2001. Land-use change model validation by a ROC (relative operating characteristic) method. Agric. Ecosyst. Environ. 85, 239–248.

Republic of Kenya, 1979. Kenya Population Census 1979, Vol. 1. Government Printer (Central Bureau of Statistics/Ministry of Economic Planning and Development), Nairobi.

Republic of Kenya, 1989. Kenya Population Census 1989, Vol. 1. Government Printer (Office of the Vice President and Ministry of Planning and National Development/Central Bureau of Statistics), Nairobi.

Riebsame, W.E., Meyer, W.B., Turner II, B.L., 1994. Modelling land use and cover as part of global environmental change. Clim. Change 28, 45–64.

Ruess, R.W., Halter, F.L., 1990. The impact of large herbivores on the Seronera woodlands, Serengeti National Park, Tanzania. Afr. J. Ecol. 28, 259–275.

Said, M.Y., Ottichilo, W.K., Sinange, R.K., Aligula, H.M., 1997. Population and distribution trends of wildlife and livestock in the Mara Ecosystem and surrounding areas: a study on the impacts of land-use on wildlife and environmental indicators in the East African Savannah. Ministry of Planning and National Development, Department of Resource Surveys and Remote Sensing, Nairobi, Kenya.

SAS Institute, 1997. SAS/STAT Software: Changes and Enhancements through Release 6.12. SAS Institute, Inc., Cary, NC.

Serneels, S., Lambin, E.F., in press. Impact of land-use changes on the wildebeest migration in the northern part of the Serengeti–Mara Ecosystem. J. Biogeogr.

Serneels, S., Said, M.Y., Lambin, E.F., 2001. Land cover changes around a major East African wildlife reserve: the Mara Ecosystem (Kenya). Int. J. Remote Sens., in press.

Singida, I., 1984. Land and population problems in Kajiado and Narok, Kenya. Afr. Stud. Rev. 27, 23–39.

Sombroek, W.G., Braun, H.M.H., van der Pouw, B.J.A., 1982. Exploratory soil map and agro-climatic zone map of Kenya, 1980, scale 1:1,000,000. Exploratory Soil Survey Report No. E1. Kenya Soil Survey, Nairobi.

Sprugel, D.G., 1991. Disturbance, equilibrium and environmental variability: what is 'natural' vegetation in a changing environment? Biol. Conserv. 58, 1–18.

Stelfox, J.G., Peden, D.G., Epp, H., Hudson, R.J., Mbugua, S.W., Agatsiva, J.L., Amuyunzu, C.L., 1986. Herbivore dynamics in southern Narok, Kenya. J. Wildl. Manage. 50, 339–347.

Stokes, M.A., Davis, C.S., Koch, G.G., 1995. Categorical Data Analysis Using the SAS System. SAS Institute, Inc., Cary, NC.

Thompson, D.M., in preparation. Livestock, cultivation and tourism, livelihood choices and conservation in Maasai Mara buffer zones, Kenya. Ph.D. Thesis. Department of Anthropology, UCL, London.

Turner, M.G., Gardner, R.H., O'Neill, R.V., 1995. Ecological dynamics at broad scales: ecosystems and landscapes. BioSci. Suppl. 1995, 29–35.

Vega Garcia, C., Woodard, P.M., Titus, S.J., Adamowicz, W.L., Lee, B.S., 1995. A logit model for predicting the daily occurrence of human caused forest fires. Int. J. Wildl. Fire 5, 101–111.

von Thünen, J.H., 1966. Der isolierte Staat in Beziehung auf Landwirtschaft un Nationalökonomie. Neudruck nach der Ausgabe letzter Hand (1842/1850). Fischer, Stuttgart.

Walsh, S.J., Evans, T.P., Welsh, W.F., Entwisle, B., Rindfuss, R.R., 1999. Scale-dependent relationships between population and environment in Northeastern Thailand. Photogramm. Eng. Remote Sens. 65, 97–105.

Wrigley, N., 1985. Categorical Data Analysis for Geographers and Environmental Scientists. Longman, New York.