Quantifying the Spatial Structure of Land Use Change: An Integrated Approach

P.H. Verburg, G.H.J. de Koning, K. Kok, A. Veldkamp*, L.O. Fresco and J. Bouma*

Wageningen Agricultural University
Agronomy/P.O. Box 341/6700 AH
Wageningen/The Netherlands
Peter.Verburg@users.agro.wau.nl
Soil Science and Geology/P.O. Box 37/
6700 AA Wageningen/The Netherlands
http://www.gis.wau.nl/~landuse1/

1. Abstract

Understanding land use change in relation to its driving factors provides essential information for land use planning and sustainable management of resources. A spatial-explicit modelling framework is proposed to study the pattern of land use and land use changes. This paper presents the methodology used to analyse the spatial pattern of land use as a function of socio-economic and biophysical factors. This analysis is used within the modelling framework to calculate the pattern of land use under scenario conditions. Results of the analysis of the land use pattern of China, Ecuador and Costa Rica are presented. These results reveal that the factors determining land use are dependent on the scale of analysis.

2. Introduction

Research has shown that human-induced conversions (e.g. deforestation) and modifications (e.g. changing land use management such as fertilizer use and irrigation practices) of land use have significance for the functioning of the earth system through their impact on biogeochemical cycles (Turner et al., 1994). Changes in these biogeochemical cycles, as a result of ecosystem conversion, can change the dynamics of greenhouse gas emissions. Land use changes can also have an important impact on the water and energy balance, directly affecting climatic conditions. The impacts of these land use changes become globally significant through their accumulative effects.

Apart from its implications for environmental sustainability, land use change can also have important consequences for food security. Conversion of cultivated land to non-farm uses such as housing, factories and infrastructure in combination with a growing population are for some countries regarded as a serious threat to future food security (e.g. Brown, 1995).
The importance of land use change for the functioning of the earth makes studies that explore land use changes very relevant. In studies on land use change it is essential to link land use changes to their driving factors. These driving factors (e.g. population or development), mediated by the socio-economic setting (e.g. market economy, resource institutions) and influenced by the existing environmental conditions or context, lead to changes in land use through the manipulation of the biophysical conditions of the land (Turner et al., 1995). Understanding trends in land use change in relation to the driving factors will provide essential information for land use planning and sustainable management of resources.
Up to now, only few models have attempted to assess the spatial patterns of the process of land use change (Hall et al., 1995). Broad scale, highly aggregated analyses of land use change at the national level provide us with information on the extent of land use change; however, it does not provide us with the spatial information needed, nor with information to calculate the consequences of land use changes. The high level of aggregation of data obscures the variability

of local situations and relationships, which can cause underestimation of the effects of land use change in certain regions and for certain groups of the population.

Spatially explicit information on land use change is needed when environmental consequences of land use changes are assessed. The amount of carbon dioxide released upon deforestation can make significant differences depending on the biomass level of primary forest (Brown et al., 1994). So, not only the amount of deforestation, but also the location and associated biomass level are required for reliable assessments of the carbon-dioxide release. Also for the calculation of changes in nutrient balances and subsequent assessments of sustainability of agro-ecosystems (Smaling and Fresco, 1993) spatial-explicit land use change data are required.

3. Socio-economic and biophysical driving factors of land use change

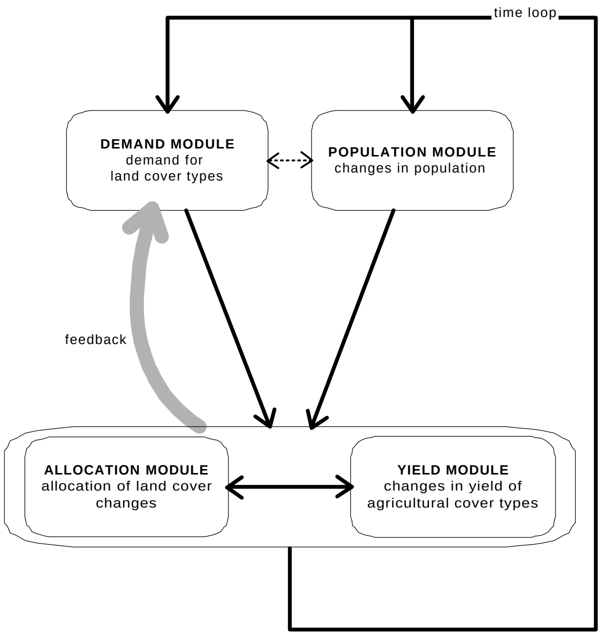
Land use changes are clearly driven by human activities (Skole and Tucker, 1993). However, the biophysical conditions of the land, such as soil characteristics, climate, topography and vegetation, determine to a large extent the spatial pattern of land use and land use change (Turner et al., 1993). An integrated approach to study the spatial pattern of land use (change) is difficult because of the complexity of the interactions between socio-economic and biophysical factors, and the different ways in which these interactions unfold in particular areas of the world. Many studies at detailed scales have been undertaken which offer detailed insights into specific cases that unfortunately cannot be generalized. Causal relations between actors and land use change, as identified on detailed scales (e.g. Groot and Kamminga, 1995), can often not be used at coarser scales. At coarse scales we need to use factors that represent or proxy the driving factors identified at the actor level. For example, to describe the process of deforestation we can use parameters like roads density, population density and average incomes as proxies for the actual drivers of land use change. Detailed studies are essential to identify which factors can be used as proxies to avoid the problem of unknown causality (does deforestation cause roads or do roads cause deforestation?).

The relations between land use and its driving factors is dependent on the scale of observation (Veldkamp and Fresco, 1996). Hall et al. (1995) found that, at detailed scales, land use in tropical rainforest areas is strongly correlated with topography. However, at a coarser scale land use was strongly correlated with precipitation and cloud cover. Most often coarse scales are useful to reveal the general trends and relations between land cover and its determining factors. Factors that influence land cover over a considerable distance (like cities) can only be observed at these coarse scales. However, the high level of aggregation at these coarse scales can obscure the variability of units and processes and is therefore considered inaccurate for detailed scale and local assessments. A multi-scale approach which identifies and quantifies the determinants of land cover at multiple scales gives the most complete description of the structure of land use.

4. Land use change modelling

To model land use changes as a function of their driving factors the CLUE modelling framework was developed. The CLUE modelling framework (The Conversion of Land Use and its Effects) is a dynamic, multi-scale, land use change model developed at Wageningen Agricultural University. The model has been used to explore land use changes in Costa Rica by Veldkamp and Fresco (1996).

The model consists of four main modules: a population module, a demand module, an allocation module and a yield module (figure 1). The demand module calculates, at the national level, changes in demand for agricultural products taking into account changes in diet and export quantity. The calculations are based upon trends of the past in combination with projections of future food demand (e.g. Islam, 1995). The population module calculates changes in population based upon different projections (e.g. Lutz et al., 1994). Historic demographic analysis is used to calculate the spatial distribution of the population. The consequences of the changes in demand and population for the pattern of land use are calculated by the allocation and yield modules. These modules take into account the variability of biophysical and socio-economic conditions within the country. On a yearly basis changes in the configuration of land use are calculated based upon an analysis of the past and present spatial distribution of land use. Feedbacks are incorporated because it is not always possible to allocate all land demanded at the national scale within the country. Such feedbacks will change the import/export quantities or lead to an adapted demand for agricultural products. The time horizon for scenario simulations is limited due to the empirical nature of the model. Simulations of reliable pathways of land use change development are not realistic for time spans longer than about 20 years.



This paper focuses on the methods used to extract the relationships between land use and biophysical and socio-economic factors. These relationships are crucial to calculate the configuration of land use in the allocation module and is therefore one of the main components of the CLUE modelling framework. The proposed methodology is applied to data from China, Ecuador and Costa Rica.

5. Data and Methods

5.1 Data

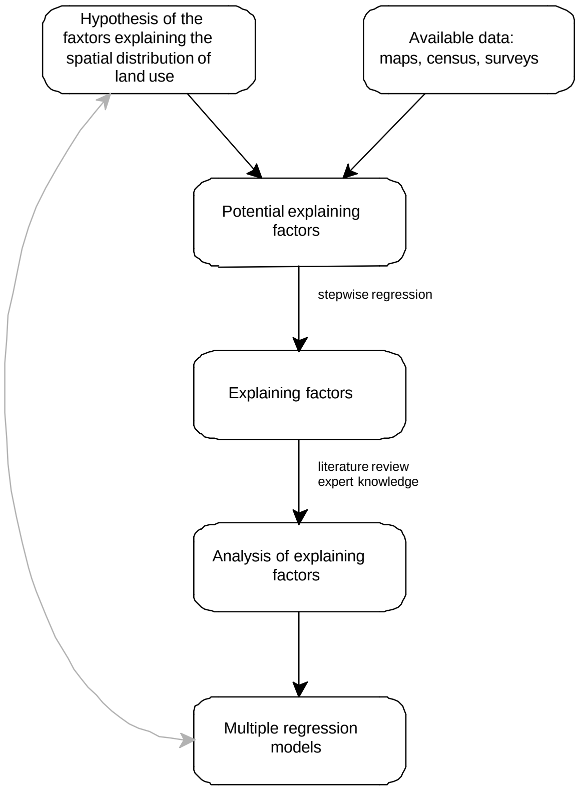
For the three different study areas biophysical and socio-economic data were collected. Potential explaining factors of the structure of land use were selected based on literature review and knowledge of the specific situation in the various countries. The data on land use and socio-economic conditions were obtained from the population and agricultural census of the countries. Census data are useful for this type of study because they contain relative extensive sets of data covering the whole country. The biophysical parameters were obtained from a number of different maps (e.g. Nuhn, 1978) or digital datasets (e.g. UNEP/DEIA, 1997), containing information on soil conditions, relief and climate. Census data can be implemented in a Geographical Information System by mapping the administrative units for which the data are derived. Because these units only rarely coincide with biophysical units a grid based system is used to facilitate analysis. Grid size was selected based on the estimated average district sizes, the most detailed spatial scale for the census data, resulting in grids of 7.5x7.5 km for Costa Rica, 9.3x9.3 km for Ecuador and 32x32 km for China.

In contrast to most grid-based approaches derived from maps which represent land use by cells with one dominant land use type, we characterize land use by the relative cover of each land use type in each grid cell, e.g. a grid cell can contain 30% cultivated land, 40% grassland and 30% forest. This way of representing the data is a direct result of the information contained in the census data.

To allow a systematic analysis of spatial scale effects, the grid data were aggregated into larger grids, composed of respectively 4, 9, 16 and 25 basic grid units, making four additional aggregated spatial scales. The new aggregated grid values were averages of the included basic grids.

5.2 Statistical methods

A stepwise regression procedure was used to identify the factors that contribute significantly (at 0.05 level) to the explanation of the land use structure. This way it was possible to distinguish which of the potential factors have relevance for the spatial pattern of land use. Expert knowledge and literature review are used to explain the relations found. Figure 2 summarizes the followed procedure. The stepwise procedure was repeated for the different aggregation levels to allow new variables to enter the regression equations at the aggregated scales. Standardized bètas were calculated to allow a comparison of the relative importance of the different identified factors.



6. Results

6.1 China

Table 1 gives the main factors that contribute to the explanation of the spatial distribution of cultivated lands and grasslands in China for the situation of 1990 (figure 3). The distribution of cultivated land can be explained reasonably well, while only 56% of the spatial variation in the distribution of grassland can be explained by a model that needs 12 different factors. The relatively poor model fit for grassland can be attributed to the wide variation in grassland types in China ranging from dry steppe in desert areas to artificial or improved grassland for livestock grazing and cutting of grass. The different types of grassland are related to different biophysical and socio-economic environments. A subdivision of grassland types would be needed to find better relations between the grassland distribution and the biophysical and socio-economic environment.

Table 1 Multiple regression models (factors and standardised regression coefficients (stb)) explaining the area devoted to cultivated land and grassland in China for 1990

cultivated land adj. r2 = 0.77		**grassland** adj. r2 = 0.56	
factor	**stb***	**factor**	**stb***
rural population density	0.56	illiterate population (%)	0.39
poorly drained soils (%)	0.19	poorly drained soils (%)	-0.35
rural population (%)	0.17	deep soils (%)	0.35
S1 soil for rainfed maize (%)[1]	0.14	temperature in coldest month	-0.27
mean altitude	-0.13	yearly precipitation	-0.24
rich fertility soils (%)	0.12	soils with low moisture storage (%)	0.24
distance to city	-0.08	moderately drained soils (%)	-0.23
		rural labour force (% of total labour force)	0.22
		rural population density	-0.18
		agricultural households (%)	-0.17
		level land (%)	0.16
		rich fertility soils (%)	0.09

*all coefficients significant at 0.05

The distribution of the rural population is the main determinant of the distribution of the cultivated lands and needs no further explanation. Cultivated lands are positively correlated with poorly drained soils while grasslands have a negative correlation with poorly and moderately drained soils. This can be explained by the relatively large share of rice farming, which requires poorly drained soils, and the vast areas of steppe occurring on well-drained soils. Cultivated land is found on soils that have a relatively high natural soil fertility and at low altitudes. The negative relation with the distance to cities can be explained by the difficulties of transporting food crops over long distances and the lack of infrastructure in more remote areas. Although climatic variables were included in the analysis they were not retained in the stepwise regression procedure, indicating that the large climatic differences in China do not influence the distribution of cultivated lands structurally. This does not imply that climatic variation does not influence the distribution of land use at all: the negative correlation with altitude is probably mainly due to the unfavourable conditions for land use in the mountains.
The distribution of grasslands, however, is strongly correlated to climatic conditions. The negative regression coefficients for temperature and precipitation indicate that grasslands are mainly found in the colder and dry regions: the northern steppe regions and grasslands on the Qinghai-Tibetan plateau. In these areas it is indeed climate that makes grassland the climax vegetation while the areas are unsuited for cultivation. The distribution of grassland is positively correlated with the percentage of the population older than 15 year that is illiterate and negatively correlated with the rural population density. We can interpret these results by considering the percentage illiterate population as a proxy of the remoteness of the area. Sparsely populated areas at large distances from the cities often have higher percentages of illiterate population. These are the areas where grasslands are found.

6.2 Ecuador

The analysis of the land use structure of Ecuador was done at 3 different aggregation levels (de Koning et al., forthcoming). In table 2, results are given for the analysis of the land use type: arable crops at three aggregation levels, giving the 7 main factors in the regression models.

The coefficient of determination increases with the aggregation level. At all levels a strong positive relation was found with rural population density. At the higher aggregation levels total

[1]Soils suitable for the cultivation of rainfed maize according to FAO, 1995. Digital Soil Map of the World and Derived Soil Properties. Version 3.5

population density becomes increasingly important, indicating the influence of urban population in the area. The positive relation with rural illiteracy (that can be considered a proxy of access to information and education) indicates the occurrence of arable crops (most of which are subsistence crops) in less endowed areas. A fairly constant negative relationship is found between arable crops and a high percentage of steep slopes, which are difficult to cultivate and have high erosion risks. Less arable crops are grown in areas with high precipitation, like the tropical climate zones in the Amazon and the northern part of the Pacific coast. Soil fertility and texture only contribute to explanation of the area arable crops at the lower aggregation levels. The preference for proximity to urban centres is likely to be related to access and size of urban markets, while rivers offer transport, access and irrigation opportunities.

It can be concluded that the factors and their relative contribution to the explanation of the area of arable crops vary with the aggregation level.

Table 2 Multiple regression models (factors and standardised regression coefficients (stb)) explaining the area arable crops in Ecuador for 1991 at different aggregation levels.

1 grid (9.3x9.3 km) $r^2 = 0.48$		9 grids (27.8x27.8 km) $r^2 = 0.65$		25 grids (46.3x46.3 km) $r^2 = 0.77$	
Factor	**stb***	**factor**	**stb***	**factor**	**stb***
rural population	0.31	rural population	0.43	rural population	0.57
Average soil fertility	-0.26	rural illiteracy	0.27	precipitation	-0.23
rural illiteracy	0.25	precipitation	-0.23	total population	0.20
Precipitation	-0.21	average soil fertility	-0.22	rural illiteracy	0.19
Sandy soils	0.20	steep slopes	-0.17	agric. labour force	0.16
Urban distance	-0.17	river distance	-0.13	steep slopes	-0.15
Steep slopes	-0.17	total population	0.10	river distance	-0.10

***all coefficients significant at 0.05**

6.3 Costa Rica

Veldkamp and Fresco (1997) did a similar type of analysis of the land use structure in Costa Rica for the situation of 1973 and 1984. At the finest level of data aggregation a stepwise regression procedure was used to identify which factors could be used to describe the land use distribution. Based on these factors multiple regression models at different aggregation levels (respectively 1, 4, 9, 16 and 25 grid cells) were calculated. This procedure keeps the same parameters in the analysis for the different aggregation levels instead of allowing new factors to enter at other aggregation levels, as was possible in the Ecuador case-study described above.
Table 3 gives the results of the multiple regression for the distribution of permanent crops in 1973. The factors that contributed to the explanation of the distribution of permanent crops were the size of the agricultural labour force, the urban population density and the amount of relief.

Table 3 Multiple regression models (factors and standardised regression coefficients) explaining the area devoted to permanent crops in Costa Rica for 1973 at different aggregation levels

Factor	1 grid (7.5x7.5 km) $r^2 = 0.45$	4 grids (15x15 km) $r^2 = 0.40$	9 grids (22.5x22.5 km) $r^2 = 0.67$	16 grids (30x30 km) $r^2 = 0.72$	25 grids (37.5x37.5 km) $r^2 = 0.66$
Agricultural labour force	0.78	0.61	0.70	0.90	0.84
Urban population	-0.18	0.11	0.22	-0.02	-0.02
Relief	-0.05	-0.04	-0.09	-0.07	-0.06

The agricultural labour force has a strong positive relationship with permanent crops at all aggregation levels, whereas urban population has alternating negative and positive relationships, and relief displays only a slight negative relationship with permanent crops. The same pattern was observed for the 1984 data (not displayed). From the data it is clear that the permanent crops are mainly found in relatively flat areas (negative relation with relief) and in areas with a substantial agricultural labour force. The changing contributions of urban population may be explained by a spatial scale effect. Permanent crops are not found too near to urban centres (negative relationship at detailed scales), but preferably at a convenient transportable distance from the urban population (positive contribution optimum at aggregation level of 9 grids).

7. Overall interpretation of the case-studies

The three presented case-studies on the spatial distribution of land use indicate that both biophysical and socio-economic factors are needed to explain the land use structure. Population density and agricultural labour force are the most important factors to explain the distribution of cultivated land. The biophysical conditions, especially soil conditions and topography, also have an important, but less pronounced, influence on the distribution of land use. Surprisingly, the influence of climate is not very structural while the climatic variations within the countries are very pronounced.

The case studies from Ecuador and Costa Rica reveal that the level of aggregation of data influences the results. This scale dependency of the land use system results in changes in the factors included in the multiple regression models (Ecuador) and in changes in the contribution of the different factors. The Costa Rica case showed that a contribution could even change from negative to positive depending on the level of aggregation. The changes with aggregation level can be attributed to changes in variability. The coarse scales are useful to identify the general trends in the land use pattern while local variability is revealed at the more detailed scales. The Costa Rica case indicated that the scale dependency is also caused by factors that act over a considerable distance such as urban population concentrations. Depending on the scale of analysis these factors will affect the relationships found.

As a consequence of the scale dependency in land use systems one should use relations derived at a certain aggregation level only for analysis and implementation at the same aggregation level. Errors will be induced when relations derived at detailed scales are used at highly aggregated scales or the other way around.

However, the scale dependencies do not only restrict our research possibilities, they also contain valuable information. By combining coarse scale analysis and detailed scale analysis a more complete description of the land use system is obtained including general trends, local variations and factors that influence land use over some distance.

8. Conclusions

This paper indicates the importance of understanding the dynamics of land use change in a spatial-explicit way. To incorporate spatial-explicit information on land use in a dynamic model the spatial structure of land use needs to be understood. A stepwise regression procedure which evaluates and quantifies the relations between potential explaining factors and land use types resulted in models that explain a considerable part of the spatial distribution of land use. Specific conditions in some parts of the countries and explaining factors not included in the analysis (like land ownership) are responsible for the remaining variability. The different case-studies indicate that both socio-economic factors and biophysical factors are needed to explain the land use structure. The relations between explaining factors and land use are scale dependent. Therefore, a multi-scale analysis is needed for a more complete description of land use. Implementation in a dynamic land use change model is useful to explore the spatial pattern of land use change under different scenarios of population growth and food demand. In combination with studies based on quantitative land evaluation methods and multi-criteria models, such as interactive linear-programming models (de Wit et al., 1988; WRR, 1992) and studies exploring yield potentials (Penning de Vries et al., 1995), this approach will lead to a more complete description of land use and land use changes.

9. Acknowledgements

This research was supported by the Dutch National Research Programme on Global Air Pollution and Climate Change (NRP).

10. References

Brown, L.R. 1995. *Who will feed china? Wake-Up call for a small planet*. The Worldwatch Environmental Alert Series. New York

Brown, S., Iverson, L.R. and A.E. Lugo. 1994. *Use of GIS for estimating potential and actual forest biomass for continental South and Southeast Asia*. In: V.H. Dale (ed.). The effect of land-use change on atmospheric CO_2 concentrations. Springer-Verlag, New York

Groot, W.T. de and E.M. Kamminga. 1995. *Forest, People, Government. A policy-oriented analysis of the social dynamics of tropical deforestation.* Main report of the project ‘Local Actors and Global Treecover Policies’. Centre of Environmental Science. CML report no. 120. Leiden

Hall, C.A.S., Tian, H., Qi, Y.,Pontius, G. And J. Cornell. 1995. *Modelling spatial and temporal patterns of tropical land use change*. Journal of Biogeography (1995) 22, 753-757

Islam, N. (ed.). 1995. Population and food in the early twenty-first century: meeting future food demands of an increasing population. International Food Policy Research Institute (IFPRI). Washington.

Koning, G.H.J. de, Veldkamp, A. and L.O. Fresco. forthcoming. Statistical analysis of land use in Ecuador at different spatial aggregation levels.

Lutz, W., Prinz, C. and J. Langgassner. 1994. *The IIASA world population scenarios to 2030.* In Lutz, W. (ed.) The Future Population of the World. What can we assume today. London.

Nuhn, H. 1978. Atlas Preliminar de Costa Rica, Informacion geografica regional. San José, Costa Rica

Penning de Vries, F.W.T., Keulen, H. van and R. Rabbinge. 1995. *Natural resources and limits of food production in 2040.* In: J. Bouma et al. (eds.), Eco-regional approaches for sustainable land use and food production, 65-87

Skole, D. and C. Tucker. 1993. Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 and 1988. Science 260: 1905-1910

Smaling, E.M.A. and L.O. Fresco. 1993. A decision-support model for monitoring nutrient balances under agricultural land use (NUTMON). Geoderma 60 (1993) 235-256

Turner II, B.L., Meyer, W.B. and D.L. Skole, 1994. *Global land use/land cover change: towards an integrated study.* AMBIO 23:91-95

Turner II, B.L., Ross, R.H. and D.L. Skole. 1993. *Relating land use and global land cover change.* IGBP report no 24, HDP report no 5., 65 pp

Turner II, B.L., Skole, D., Sanderson, S., Fischer, G., Fresco, L. and R. Leemans. 1995. *Land-Use and Land-Cover Change; Science/Research Plan*. IGBP Report no. 35; HDP Report No. 7

UNEP/DEIA (1997). Singh, A., Fosnight, E. and Rykhus, R. *Spatial Data Sets for Environmental Assessment: Towards Bridging the Data Gap* - UNEP/DEIA/TR.97-4. Division of Environment Information and Assessment. United Nations Environment Programme

Veldkamp, A. and L.O. Fresco, 1996. CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. Ecological Modelling 91: 231-248

Veldkamp, A. and L.O. Fresco, 1997. Reconstructing Land Use Drivers and their Spatial Scale Dependence for Costa Rica (1973 and 1984). Agricultural Systems, Vol. 55, No. 1, pp. 19-43

Wit, C.T. de, Keulen, H. van, Seligman, N.G. and I. Spharim. 1988. Application of interactive multiple goal programming techniques for analysis and planning of regional agricultural development. Agricultural Systems 26: 211-230

WRR (Netherlands Scientific council for Government Policy). 1992. *Ground for choices. Four perspectives for the rural areas in the European Community*. Report to the Government 42, SDU Uitgeverij, The Hague, 144 pp