Modelling to explore land use patterns at the forest edge: Objectives and Model Design

Jerome K. Vanclay
Center for International Forestry Research, Bogor, Indonesia

Abstract A model is proposed that will enable forest and land use policy options to be explored in an repeatable, objective and quantitative way. Construction and use of the model should lead to a better understanding of spatial and temporal land use patterns. It will provide detailed spatially-explicit data on a range of parameters to enable on-going development and validation. The model is based on the assumption that land use patterns ultimately are shaped by individuals who make rational decisions based on available information, obligations and expectations (social as well as economic). Since these decisions are rational, they, and the resulting land use patterns, can be modelled and predicted. Furthermore, we assume that in making these decisions, individuals balance the anticipated returns and risks, and that an individual's particular strategy (i.e. emphasis on profit-maximization or risk minimization) reflects his status and security. Important inputs to the model include spatial (land tenure, topography, soils), social (demography, clan obligations, cultural traditions), and economic data (crop yields, market prices, transport costs). Notice that land use is not an input to the model, but that the model can predict spatially-explicit land use patterns. Thus land use predictions may provide a robust way to benchmark the model. The model can be implemented in a hierarchical way, so that it can operate at the village, provincial or national scale, with the detail of inputs and outputs varying accordingly.

1. INTRODUCTION

One of the weaknesses of much forest policy research is the difficulty of demonstrating robust empirical tests of the propositions put forward. One way to allow more robust testing is by modelling the social and ecological factors of interest, and comparing model predictions with empirical data. Although such models seem feasible, no spatially-explicit models of this kind appear to have been implemented to date. Models are not the only way to test propositions, but a major attraction of formal modelling is that ideas must be expressed completely, concisely and explicitly, and implemented in an integrated and testable way. They also offer new insights and pose new problems for research: "More information can be read from a map than was needed to construct it" [Ziman 1978].

In this paper, I discuss the rationale for a model designed to improve understanding of land use patterns in time and space, especially near the forest edge, and to help us explore in a quantitative way, policy options intended to manipulate these patterns. It is inevitable that initial attempts to construct models of this kind will at first be simplistic. However, simple models may still offer powerful insights. In particular, models excel at exposing counter-intuitive consequences of simple assumptions. Even if initial prototypes of the model are of little practical relevance, these prototypes may offer valuable insights, and their main purpose may be to sharpen questions rather than to provide answers.

The basic concepts in this work are not new; what is new is the way concepts are integrated and applied. Much of this work develops from the basis established by von Thünen [1826]. Some recent work [e.g. Dunning et al 1995, Flechsig et al 1994, Wilkie and Finn 1988, Lambin 1994] touches on the concepts expounded here, but no-one seems to have explored the particular implementation proposed here. Other models [e.g. de Klein 1989] exploring a similar topic using analogous approaches (viz. system-dynamics) have tended to employ rather simplistic relationships, and have not attempted spatially-explicit predictions.

Testing remains one of the weak points in much modelling work [e.g. Meadows and Robinson 1985, Vanclay et al 1995]. There is no substitute for formal empirical tests, as subjective appraisals (even comparatively straight forward appraisals of technical equipment) may be misleading [cf. Skovsgaard et al. 1995].

The proposed model should provide detailed spatially-explicit data on a range of parameters to enable on-going model development and testing. Map-based summaries may be a particularly useful form of output that is easily interpreted and tested. However, map-based output is feasible only at the village scale, and is impractical at the national scale, so aggregate indices such as social indicators of well-being and equality, and ecological indicators of biodiversity and sustainability should also be provided.

2. ASSUMPTIONS

The proposed model relies on four basic assumptions, namely that:

1. Land use patterns are ultimately shaped by individuals and groups of individuals;

2. These individuals make *rational* decisions based on available information, obligations and expectations (social as well as economic; note that perceptions may be more important than reality);
3. Individuals tend to maximize expected benefits or to minimize anticipated risks;
4. Both benefit-seeking and risk-avoidance can be modelled by maximizing the risk-adjusted benefits (see below).

Decisions affecting land use patterns may typically involve the production of one or more products to achieve the maximum benefit subject to some social and economic constraints. Note that risk-avoidance may be modelled by introducing a suitable discount factor to account for the anticipated risk. Thus the benefit to an individual k may be estimated by choosing products i and sites j so as to maximize

$$\Sigma_j \{Max_i\ [Yield_{ij} \times Price_{ij} \times Risk_{ijk} \times Share_{ijk} - Input_{ij} - Sell_{ij}\ ]\} \quad (1)$$

Or in non-mathematical terms, choose the "best" combination of products for each of the sites available to the individual, so that the overall benefit to the individual is maximized. Note that "best" depends on many things: the *anticipated* yield for that activity (e.g. crop, handicraft item, wage-based employment, etc) at that site, the anticipated price, any reduction for real or imagined risks (pests, disease, fire, theft, loss of tenure, spoiling during transport to market, etc), an allowance for shares that others may have in the activity (clan obligations as well as landlords who may share revenues but not costs). For efficiency, we assume that yields and prices are the same for all individuals (ignoring production and negotiating skills), but recognize that individuals may differ in their willingness to accept risks, and in their social obligations. Notice that the correction for risk ($Risk_{ijk}$) in equation 1 may reflect the long-run expectation for individuals who are not risk-adverse, but may be substantially less for those who are unable or unwilling to contemplate a risky venture. The gross return to the individual has to be adjusted for the costs of production ($Input_{ij}$) and the costs of marketing ($Sell_{ij}$). Note that these also depend on the product and the site. Production costs may include labour (own or paid), rent (formal or informal obligations), and other inputs (fertilizers, pesticides, etc). Individuals may undertake activities because of the status conferred, and the production cost may need to be adjusted for this perceived status. Marketing costs include transport and packing, and in some situations, advertising.

This is not a typical linear programming problem formulation, as each individual may attempt to find an optimal solution for their family or clan, even if it leads to a sub-optimal outcome for the village as a whole [cf. Hardin 1968].

The terms in equation (1) can all be quantified in some sense, so it should be possible to construct and evaluate this model, provided that we can quantify them all in similar units (e.g. dollars or other local currency). Note that history does not enter this equation, except in that it influences the choice of activities i entertained by each individual. This could be formulated as an interesting mathematical programming exercise, but it is probably sufficient (at least initially) to solve it heuristically (i.e. trial and error with a simple set of guidelines).

The decision made for any particular site j is not independent of decisions made for other sites; price and risk depend on total production across all sites - otherwise everyone might grow the same high-value crop, saturating the market. Thus the model should account for the potential demand and price elasticity at each regional market centre. It might also include lagged adjustments to take into account the time taken to learn and implement new technologies. However, in an initial prototype of the model, we could avoid this complexity by making the prevailing market prices *external* to the model, getting the user to provide the prices and assuming that they remain constant. This avoids many complexities, since the actual prices paid may depend on elasticities, the number of producers and buyers, and local wealth (substituting luxury for inferior goods with increasing wealth). It also simplifies the model since we can then assume decisions on any site are independent of other sites, so that equation (1) can be solved without taking time into account (i.e. a *static spatial* model, easily implemented as a geographic information system).

In the simplest case, we can imagine equation (1) applying to agricultural monocultures, but conceptually, it is possible to apply the same approach to mixed plantings (e.g. agroforestry; but this makes yield tables complicated) and to alternative employment opportunities (including industrial).

3. INITIAL IMPLEMENTATION

As outlined above, the model could be used in a mathematical way to evaluate land use options for a few sites, but to put it into a useful context, it should be implemented within a geographic information system (GIS) framework. If we provide one or two elements of the equation as external inputs (e.g. prices), then the equation can be deterministic, without reference to time. Then all the details needed to evaluate decisions can be taken from data normally available in a GIS, e.g.

Predict	*from*
Yields	soils, topography, climate, historic data
Prices	external data provided for each product
Risk	tenure, topography, socio-political & historic data
Share	tenure, social data (ie family/clan obligations)
Inputs	soils, existing vegetation, capital, cultural traditions
Sell	distance, infrastructure, social data

Many of these raw data requirements can be obtained from a conventional GIS, which typically may provide information such as soils, topography (i.e. digital elevation model), climate, vegetation, tenure, transport routes and urban areas. Crop yields may be estimated from historic data, standard yield tables, or from software such as Plantgro [Hackett 1991]. Costs of marketing may be estimated from the time and energy expended in transporting the product by road, rail, river and beasts of burden.

Since land use does not enter, but can be predicted by, equation (1), the predicted land use patterns may provide a robust test of the model. However, some care may be required in interpreting such comparisons, and may be appropriate to amalgamate some activities into broad classes.

This sort of static spatial model will provide a "picture" that will respond to changes in input parameters, but predicted land use classes for much of the study area may remain fairly static, despite moderate perturbations in input variables. We want to identify the "sensitive" areas, where comparatively small perturbations give rise to large changes in predictions. In particular, we want to know where these areas are, what parameters trigger shifts in dominant land use, and how these shifts occur. Initial preconceptions suggest that these sensitive areas may be near the forest edge, and may include *imperata* grasslands. However, to establish or refute this, we need to do sensitivity analyses on all input parameters.

4. FURTHER DEVELOPMENTS

If conditions change so that land use becomes less intensive, natural vegetation may not reappear immediately, but may take time to regenerate. This has implications for other land uses, as the nature of existing vegetation may influence land use decisions (e.g. cost of clearing or weeding; potential to harvest established plants/trees). This "fallow" or regenerating "abandoned" land may be best modelled dynamically. It seems inefficient to dynamically model every cell in the GIS, so selected sites may form the basis for interpolation. Techniques such as adaptive simulated annealing may provide a near-optimal basis for sampling [Ingber 1993].

We may also need dynamic models to be able to investigate if a given level of harvesting (of timber or non-timber products) in forested areas is sustainable, or if not, to calculate the time to depletion. A further challenge for later versions is to model species interactions, especially for apparently pivotal (keystone) species, as these may influence regrowth on "abandoned" land.

The proposal outlined above is easy to conceive for a small village, where we can include every individual in the solution to equation (1). However, if we scale up to include a bigger population, it becomes clear that we cannot examine decisions individual-by-individual, and must extrapolate from a sample of individuals. The choice of sample may be critical to the outcome, and suitable sampling strategies should be used. However, the approach can be scaled-up to the local, provincial or national level, provided that sufficient data are available and a suitable sampling strategy is employed. A crucial part of this is to identify the minimum essential set of prime determinants. If we can agree that land use is shaped primarily by one or two factors, e.g. transport costs and alternative employment opportunities, then we will be able to construct an efficient model at the national level using data drawn largely from GIS systems. However, as the number of critical determinants increases, the computational load and data requirement increases rapidly, and it may not be feasible to scale-up a complex detailed model. Hence simplicity is a great advantage in applying the model.

There are several other specific problems that need to be addressed before this model can be realized in its proposed form. Some anticipated problems include the implementation of efficient travel cost/time algorithms based on GIS data [e.g. Wilkie and Finn 1988], and efficient ways to obtain reliable site-specific yields, inputs and prices for all potential crops and activities. Although some algorithms are currently available, the proposed model will require efficient generalizations of current methods.

5. PROGNOSIS AND POTENTIAL IMPLICATIONS

Is the proposed model worth attempting? I make no attempt to conceal the difficulties of constructing and implementing a useful version of this model, but there are many benefits, including the knowledge and experience that will be learned during the attempt. The model will probably be completed in an iterative way, with many prototypes before a useful model is attained. Each prototype should offer the opportunity to test and refine ideas a little further, to advance out understanding and knowledge, and to smooth the path for subsequent attempts. When a useful model is completed, it should revolutionize the way we explore forest policy options and land use implications.

6. ACKNOWLEDGMENTS

Many colleagues have shared helpful insights which have influenced the present proposal. I am especially grateful to Neil Byron for his guidance.

7. REFERENCES

Dunning, J.B. Jr., D.J. Stewart, B.J. Danielson, B.R. Noon, T.L. Root, R.H. Lamberson and E.E. Stevens, Spatially explicit population models: current forms and future uses. *Ecological Applications* 5, 3-11, 1995.

Flechsig, M., M. Erhard and V. Wenzel, Simulation-based regional models — concept, design and application. *Ecological Modelling* 75/76, 601-608, 1994.

Hackett, C., Plantgro: a software package for coarse prediction of plant growth. CSIRO, Melbourne. 1991.

Hardin, G.J., The tragedy of the commons. *Science* 162, 1243-1248, 1968.

Ingber, A.L., Simulated annealing: Practice versus theory. *J Mathl. Comput. Modelling* 18, 29-57, 1993.

de Klein, J.P.G. Analysis of land use development: a system dynamics land use model for Gunung Kidul (Java). Fakultas Kehutanan, Universitas Gadjah Mada, FONC Project Communication No 1989-5, 78 pp., 1989.

Lambin, E.F., Modelling deforestation processes: a review. TREES Series B Research Report No 1 (EUR 15744), Luxembourg, 1994.

Meadows, D.H. and Robinson, J.M., The *Electronic Oracle: computer models and social decisions.* Wiley, xv+445 pp., Chichester, 1985.

Skovsgaard, J.P., V.K. Johannsen and J.K. Vanclay, Laser dendrometers in forestry: tools or toys? in *Applications of New Technologies in Forestry*, edited by L. Bren and C. Greenwood, pp. 289-298, Institute of Foresters of Australia, Canberra, 1995.

Thünen, J.H. von, *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*, 1826. English translation by P. Hall *Von Thünen's Isolated State. Pergamon Press*, 1966.

Vanclay, J.K., J.P. Skovsgaard and O. Garcia, Evaluating forest growth models, paper presented at International Union of Forestry Research Organizations (IUFRO) XX Congress, Tampere, Finland, 6-12 August, 1995.

Wilkie, D.S. and J.T. Finn, A spatial model of land use and forest regeneration in the Ituri forest of northeastern Zaire. *Ecological Modelling* 41, 307-323, 1988.

Ziman, J., *Reliable Knowledge: an exploration of the grounds for belief in science*. Cambridge Univ. Press, 197 pp., Cambridge, 1978.