Available online at www.sciencedirect.com
SCIENCE DIRECT®
Ecological Modelling 185 (2005) 105–131
ECOLOGICAL MODELLING
www.elsevier.com/locate/ecolmodel

Tropical deforestation in Madagascar: analysis using hierarchical, spatially explicit, Bayesian regression models

Deepak K. Agarwal[a,*], John A. Silander Jr.[b], Alan E. Gelfand[c], Robert E. Dewar[d], John G. Mickelson Jr.[e]

[a] *AT&T Shannon Research Labs, Florham Park, NJ 07932-0791, USA*
[b] *Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269-3043, USA*
[c] *Department of Statistics and Decision Sciences, Duke University, Durham, NC 27708-0251, USA*
[d] *Department of Anthropology, University of Connecticut, Storrs, CT 062692176, USA*
[e] *Center for International Earth Science Information Network, Columbia University, Palisades, NY 10964, USA*

Received 4 August 2003; received in revised form 17 August 2004; accepted 29 November 2004

Abstract

Establishing cause–effect relationships for deforestation at various scales has proven difficult even when rates of deforestation appear well documented. There is a need for better explanatory models, which also provide insight into the process of deforestation. We propose a novel hierarchical modeling specification incorporating spatial association. The hierarchical aspect allows us to accommodate misalignment between the land-use (response) data layer and explanatory data layers. Spatial structure seems appropriate due to the inherently spatial nature of land use and data layers explaining land use. Typically, there will be missing values or holes in the response data. To accommodate this we propose an imputation strategy. We apply our modeling approach to develop a novel deforestation model for the eastern wet forested zone of Madagascar, a global rain forest "hot spot". Using five data layers created for this region, we fit a suitable spatial hierarchical model. Though fitting such models is computationally much more demanding than fitting more standard models, we show that the resulting interpretation is much richer. Also, we employ a model choice criterion to argue that our fully Bayesian model performs better than simpler ones. To the best of our knowledge, this is the first work that applies hierarchical Bayesian modeling techniques to study deforestation processes. We conclude with a discussion of our findings and an indication of the broader ecological applicability of our modeling style.

Keywords: Tropical deforestation; Bayesian statistical hierarchical models; Spatially explicit regression models; Misaligned spatial data; Human population pressure; Land-use classification; Statistical model comparison; Multiple imputation

1. Introduction

The demise of the world's tropical rain forests has been of central concern to conservation biologists for

* Corresponding author. Tel.: +1 9733607154; fax: +1 9733608178.
E-mail address: dagarwal@research.att.com (D.K. Agarwal).

0304-3800/$ – see front matter
doi:10.1016/j.ecolmodel.2004.11.023

at least 20 years (Singh, 2001; Whitmore, 1998). But cognizance of the problem was apparent decades earlier (cf. Jarosz, 1993; Perrier de la Bâthie, 1921). Yet it has proven surprisingly difficult and controversial to obtain accurate estimates of deforestation rates or indeed to relate current forest cover to potential or historical extent of forest cover. This remains true even with the advent of satellite imagery from the 1970s (Apan, 1999). The difficulties arise from issues of definition and benchmarks, spatial and temporal resolution, and data adequacy (Singh, 2001; Angelsen and Kaimowitz, 1999). As a consequence even the most recent estimates of deforestation rates, or more generally land-use change, vary considerably, and debate continues over whether the rates of deforestation in the tropics are declining or increasing (FAO, 2001; Matthews et al., 2000). Only for specific, narrowly defined regions can one obtain reasonably reliable data on deforestation rates over specific time periods (e.g., Mertens and Lambin, 2000).

Establishing cause–effect relationships for deforestation or land use at various scales has also proven difficult even when the process appears well documented (Angelsen and Kaimowitz, 1999; Irwin and Geoghegan, 2001; Barbier, 2001). The standard explanation for deforestation in the tropics has been rapid population growth, associated poverty, and consequential environmental destruction (Leach and Mearns, 1988; Richards and Tucker, 1988; Mercier, 1991; Brown and Pearce, 1994; Sponsel et al., 1996). In selected areas, commercial exploitation and clear cutting are also important causes of deforestation (Torsten, 1992). However, conventional explanations for forest loss and environmental degradation have recently been questioned as too simplistic, general, or even misleading (Barbier, 2001). In the absence of a clear understanding of the role that various contributing explanatory variables may play in deforestation processes, let alone establishing cause effect relationships, it is not surprising that reforestation or forest conservation schemes have had relatively little or no success. Most of the reports on failures are buried in the gray or secondary literature (e.g., Bloom, 1998; IUFRO, 2001; Sharma et al., 1994; cf. Oates, 1999); relatively few are to be found in the primary literature (e.g., Olson, 1984; Elster, 2000); and rarely does one find reports of successful reforestation either via natural succession (e.g., Guariguata and Ostertag, 2001) or by human management practices (e.g., Lamb et al., 1997).

In this paper, we focus on deforestation processes and patterns of land use in the eastern wet tropical forests of Madagascar (Fig. 1). The prevailing perception is that the patterns and processes here are well understood. In fact the land-use change maps of Green and Sussman (1990) are commonly reproduced in textbooks as standard examples of deforestation processes. Nevertheless current and historical patterns of deforestation in Madagascar remain poorly understood, and controversial (e.g., Jarosz, 1993 versus Green and Sussman, 1990). Current estimates of forest cover in Madagascar vary by at least five-fold, from 42,000 km^2 for "closed forest," up to 158,000 km^2 for "natural forest," to 232,000 km^2 for "total forest and woodland," respectively, about 7, 27 and 40% of the land area of Madagascar (UNEP, 1998; UNEP-WCMC, 2001). This underscores the pervasive difficulty of estimating forest cover, which in part reflects differing definitions of forest (cf. Silander, 2000). In contrast, the forest cover in 1902 was estimated at about 20% of the land area (Pelet, 1902). In 1921, based on extensive field reconnaissance from 1900 to 1915, Perrier de la Bâthie estimated that about 19% (110,000 km^2) of Madagascar was in forest or woodland of one sort or another. But of this only about 35–70,000 km^2 (6–12%) was closed forest. In 1934, forest cover was estimated at only 10% of the land area (Grandidier, 1934). These figures conflict with estimates of forest cover in the 1960s: 33% forested in the early to mid 1960s (Humbert and Cours Darne, 1965) versus 21% a few years later (Le Bourdiec et al., 1969). The fact that these contemporary estimates of forest cover are so different is surprising, given the availability of aerial photography and more precise cartographic methodology. Satellite data from the 1970s onward have been able to provide only limited estimates of forest cover or deforestation rates. Virtually all of the available images of the wet tropical forests of the eastern Madagascar were too obscured by clouds through the 1990s to provide adequate estimates of forest cover, especially for the focal region of this study. In fact, the most recent available maps reproduce "current" forest cover for most of the east coast from the maps published in 1965 (Faramalala, 1995; Du Puy and Moat, 1999; see also Green and Sussman, 1990) with no new data. These sources thus indicate current forest cover for our focal region (Taomasina Province)

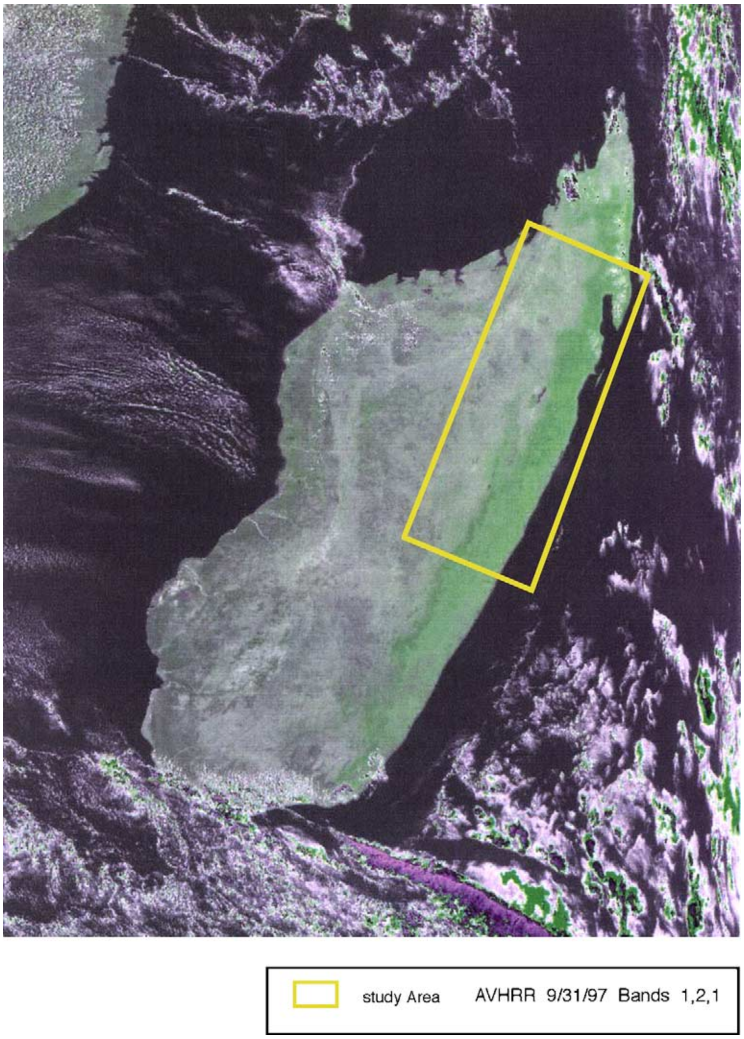

Fig. 1. Color composite AVHRR satellite image of Madagascar. The box encompassing Taornasina Province shows the focal area for the study. The light and dark green areas in along the east coats of Madagascar correspond roughly to potential forest and mature, continuous forest, respectively.

at 66%, while the center of our area (Fenerive Prefecture) would be 95% forest cover. Clearly this is a gross overestimate of actual forest cover (see results below), and a marked departure from all earlier estimates in the twentieth century, and is indeed at odds with historical accounts from the 17th through 19th centuries. Even Green and Sussman's seminal 1990 study of deforestation rates in Madagascar are based largely on the same set of maps provided by Humbert et al. (1964–1965).

The confusion and contradiction described above, underscores the need for a more critical, in depth examination of deforestation or more generally land-use patterns and processes. It also points up the need for evaluating land-use patterns in a more thorough historiography context. Moreover, if there is any indication that patterns of land-use change spatially from one region to another, this needs to be addressed or one may risk obscuring pattern and process across scales.

With these objectives in mind, we present a novel hierarchical Bayesian model for explaining deforestation processes applied to deforestation of Madagascar. Our method is able to address several issues that arise in the study of deforestation processes, which, to the best of our knowledge, has not been addressed before. First, we provide a framework to connect data layers that are spatially misaligned. In our example, we explain land-use patterns (measured at a resolution of $1\,\text{km} \times 1\,\text{km}$) in terms of population counts available at a much coarser resolution (by townships). We show that naive schemes like allocating population uniformly to all pixels within a given township perform poorly compared to the multilevel hierarchical Bayesian model we suggest. Next, we provide a novel multiple imputation method to deal with missing data when modeling spatial processes like deforestation. Note that "filling up holes" created due to missing land-use classification using a crude method like "majority vote due to neighbors" might fail to provide correct estimates of uncertainty for model parameters. Finally, our model can explicitly incorporate spatial structure, which has hitherto not been addressed in the context of explanatory models that study deforestation processes. Apart from providing better statistical inference by modeling excess heterogeneity present in the data, maps of estimated spatial structure are informative and can often lead to new ecological insights. However, this comes at a price of being able to fit complex hierarchical Bayesian models, which is a computationally challenging undertaking and may involve writing special purpose code as is the case here. Hierarchical modeling is just beginning to receive attention in the ecology literature (see our brief review of this work at the end of Section 2). Beyond the obvious applications to land change problems, hierarchical modeling has broad application to other ecological patterns and processes which present many of the same issues listed above. For example, in attempting to explain species presence–absence, abundance or richness using ecological explanatory variables, such as climate, soil geology or topography, one must address missing data, spatially misaligned explanatory data layers, and spatial structure in the ecological pattern. Because we had some a priori indication that there were sub-regional differences in land use within our study region, we subdivided the analysis spatially into separate northern and southern sub-regions. Finally we interpret the results of the model in the context of the critical historical information available. Because the modeling approach we take is novel and technically complex, we start with detailed discussion of the modeling approaches.

2. Model background and modeling issues

The goal of this paper is thus to develop, fit, and interpret stochastic models to clarify spatial processes that can explain patterns of deforestation or more generally land-use patterns. This is a rather challenging undertaking on several accounts.

First, there are numerous factors which have been linked to land use and deforestation. These factors can be socio-economic, e.g., population growth or economic growth; physical factors, e.g., topography or proximity of rivers and roads; policy-driven government intervention, e.g., agriculture and/or forestry policies; external factors, e.g., demand for exports or financing conditions. For any given region, typically only some of this information is available or selectively available at different spatial scales.

Second, there is no well-accepted notion of a response variable. In some cases deforestation rates or forest areas may be used. Deforestation rates are typically calculated over limited time spans. Forested areas are usually obtained from cross-sectional studies of different countries or different regions within countries. The alternative, which we adopt is to partition the study region into disjoint areal units and then attach an ordinal land-use classification variable to each unit. This classification represents successive departure from a potential, fully forested landscape, which all evidence points to for eastern Madagascar. Of course, such a variable is not uniquely defined in terms of number of and definition of classifications.

Third, the explanatory variables and the response variables are typically measured in different areal units. For instance, in the dataset we investigate, the response variable is land-use classification (e.g., forested, non-forested, etc.), which is ascribed to $1\,\text{km} \times 1\,\text{km}$ pixels derived from satellite images. On the other hand, population is recorded at various administrative areal units. In our case, we use the equivalent of "township" level data, which are considered to be the most reliable census data at the finest spatial resolution. The resultant challenge is to develop a regression model for data collected on spatially misaligned areal units, i.e., areal units (pixels, polygons, etc.) that do not correspond to each other or line up as GIS data layers.

Finally, land use and deforestation are inherently spatial processes. So too, are many of the explanatory variables, such as population counts. Satisfying stochastic, as well as mechanistic objectives, modeling should capture association between measurements on areal units in terms of proximity of these units. One way to accomplish this is to introduce appropriate sets of spatial random effects, which can serve as spatial surrogates for unmeasured or unavailable covariates that are inevitable in any modeling protocol.

Little formal modeling of deforestation was attempted until the 1980s (Granger, 1998). Descriptive statistical summaries of land use or forest area obtained for certain regions at certain time points were customary. Most of the ensuing statistical work, which has appeared has been based on standard multiple linear regression models relating deforestation rates or forest area to a long list of potential explanatory variables (Granger, 1998). In fact, Granger lists at least 28 different variables, which have been linked directly or indirectly to deforestation or land-use change in a forested landscape, while Anglesen and Kaimowitz list 140 different models developed to explain just economic causes of deforestation. In these studies it appears that the major objective has been simply to maximize R^2; and models with eight or more explanatory variables have been put forward. This is no longer considered sufficient or useful in statistical modeling these days. Moreover, none of these models are explicitly spatial in nature. The little work with an explicitly spatial flavor, which does exist has arisen from an econometric perspective introduced by Palloni (1992) and Chomitz and Gray (1996), and more recently summarized by Irwin and Geoghegan (2001). These models assume that land use will be devoted to the activity yielding the highest "rent". The modeling connects output prices to input prices through "production" functions with the spatial aspect introduced solely by relating these prices to market distance. These models remain quasi-spatial and econometric.

Hence, we claim that despite the already substantial literature, there is a need for better, explanatory models, which may provide better insight into the process of deforestation or land-use change (cf. Irwin and Geoghegan, 2001; Veldkamp and Lambin, 2001). Various modeling lineages have been developed and applied to explain deforestation with questionable success and little attempt at synthesis. These include deterministic and stochastic models, statistical and simulation models, phenomenological and mechanistic models, spatial and non-spatial models, etc.

In general, stochastic models provide considerable advantage over deterministic approaches in situations where substantial variability is always present, as in land-use change data. Stochastic modeling allows one to resolve the signal and to clarify the nature of the noise, which obscures it. The development of satisfying stochastic models does raise a number of methodological issues. Here, we offer an overview of a critical subset of these issues. We address the specific details of these issues later as part of the development section.

First, there is a matter of modeling style. Do we employ simulation modeling, say as is frequently done with cellular automata models, attempting to capture mechanistic aspects of the process? Or do we employ phenomenological modeling to attempt to relate different data layers, each measuring different variables? Where there is theory to guide us, can we be somewhat mechanistic in relating data layers? Finally, do we take a static or dynamic view of the process? There is no simple right or wrong answer to these choices.

If we seek to explain land use, then we first have to formally define the response variable. Even within a specific context there is not usually a well-accepted notion. Also, we have to identify which factors we seek to link to land-use. Do we seek a socio-economic explanation? Do we seek a physical or ecological explanation? Here we must recognize that no proposed explanatory model is *correct*, but that some are more *useful* than others and that no proposed explanatory model establishes causality, merely relationship. Moreover, models employing different subsets of variables can explain comparably well. One must decide which perspective to explore and which subset of possible factors to connect.

When data layers are to be interrelated, they can be introduced at various levels of a hierarchical specification, which becomes an explanatory model for the process (cf. Wu and David, 2002). In this specification the lowest level is the response. This modeling strategy is referred to as hierarchical or multilevel modeling. Here, for example, the hierarchical levels of the full model include a population model, a land-use classfication model, and a model specifying latent variables. This approach allows for considerable flexibility in modeling and the possibility of incorporating mechanistic components in linking certain levels. Hierarchical modeling is achieving increased utilization in analyzing data collected from complex processes. See the recent papers by Wikle (2003), by Clark et al. (2004), and by Gelfand et al. (2003, 2005).

Full inference for hierarchical models is most easily achieved (and often only possible) within a Bayesian framework. In particular, this framework enables a posterior distribution (i.e., a conditional distribution given the observed data) for all model unknowns. This distribution allows any desired inference about any of the unknowns. Such models are fitted using simulation-based methods, e.g., Gibbs sampling and Markov chain Monte-Carlo methods. See Carlin and Louis (2000) and Gelman et al. (1995) for an introductory presentation of this methodology.

In the context of explaining land use, the data layers are inherently spatial. Some are observed at particular point locations, and some are obtained through satellite imagery as raster pixel values at some degree of resolution. Other data may be associated with formal governmental units such as towns or other municipal subdivisions, stored as polygons. Data, which is obviously spatial should be modeled in an explicitly spatial fashion. Only rarely has this actually been done in modeling land-use patterns, and where spatial models have been developed, the analysis is not entirely spatially continuous (Irwin and Geoghegan, 2001). Developing spatial process models that lack spatial contiguity is analogous to developing models of time series processes that lack knowledge of the time order of the processes—not very useful (Irwin and Geoghegan, 2001). While there may be issues of model choice in providing spatial structure as well as the scale or resolution at which the modeling is done, it seems clear that measurements which are closer to each other in space should be more strongly associated than those farther from each other. Spatial dependence should be incorporated, but since different spatial data layers are often collected at different scales, one needs a formal (rather than ad hoc) modeling strategy to carry out full inference in the presence of such consequential misalignment. Bayesian hierarchical modeling provides this framework.

In this regard, some spatial features can be described through the mean structure of the model (or in the case of categorical or count data, the mean on a transformed scale). For instance, one can introduce distance to a road or to a town center with a coefficient as part of a regression structure. Or one can attempt to explain spatial features through a trend surface in the mean (e.g., a polynomial of low order in latitude and longitude). But there will always be omitted, perhaps unobservable, explanatory variables, which also provide spatial explanation. In the absence of these variables, spatial association in the residuals will be expected, and hence models introducing spatial dependence are needed.

For at least some data layers there will be missing data. For example, land-use classifications obtained by remotely sensed satellite data, particularly for wet tropical regions, will have portions obscured (hence missing) by cloud cover. Typically missing data are treated in an ad hoc manner. Models of spatial association for gridded data assume there are no missing observations. To accommodate such *holes* in the dataset, imputation (i.e., statistical simulation of missing values) provides an alternative to subjective assignments of values. That is to say, one imputes observations at the missing locations according to the likelihood of the different possible values. Such imputation must be done randomly and, moreover, *multiple*, hence variable, imputations are necessary in order to capture the effect of the uncertainty associated with imputation and to better reflect variability associated with inference in the presence of missing values.

Lastly, the multilevel hierarchical approach allows for the specification of many different models; there is flexibility at each level. But, with many possible models to consider, tools for model comparison are required. Simple notions like R^2 are clearly inappropriate for multilevel models with categorical or count variables as response. Criteria, which reflect the utility for the model emerge as more attractive and better suited. However, such comparison is really only sensible across models employing the same response data and with explanatory objectives with regard to available data layers. So, these are the only model comparisons we can provide.

Hierarchical Bayes models have a long and successful history by now in several areas but are just beginning to be applied to problems in ecology (e.g., Borsuk et al., 2004; Qian et al., 2003; Borsuk et al., 2001; Prato, 2000; Clark, 2003). See Banerjee et al. (2003) for a general introduction to the methodology. However, to the best of our knowledge, the methodology has not been applied to study deforestation processes other than a preliminary paper we have published in the statistical literature (Agarwal et al., 2002). Before providing details on our model development and specification in next section, we would like to comment briefly on the difference between our approach and the simulation modeling approach that has been widely used to study deforestation (see, for example, Soares-Filho et al., 2002; Duffy et al., 2001; Kok and Winograd, 2002). We remark the latter is really a different philosophic approach to the problem compared to the explanatory approach we adopt. While simulation-based modeling seeks to predict observed land-use behavior by capturing the known mechanistic aspects of the problem, our explanatory approach attempts to explain a response variable (in this case land use) in terms of other explanatory variables. Moreover, simulation models are dynamic, seeking to compare the evolution of deforestation. Our model is static seeking to link land use to explanatory variables, in particular to population pressure. As such, the two approaches are not directly comparable to understanding deforestation patterns and processes. In fact, the two approaches can in some cases complement each other—insights into the mechanistic aspects of the problems can be incorporated when building an explanatory model. However, this has rarely been attempted.

3. Overview of model development and specification

Our approach is to formulate a model for the joint distribution of local human population attributes and forest exploitation, given other explanatory variables that are explicitly spatial. In so doing we are immediately faced with incompatibility of the data layers (land use is obtained at the scale of 1 km × 1 km pixels (see below) while population counts are obtained at irregular town-level polygons). We then overlay the town-level map on the pixel-level map and modify town boundaries, using *majority pixel rule*, so that each pixel is contained in one and only one town. Upon such *rasterization* of the town data, we conduct the joint modeling at the pixel level. This joint distribution is specified by modeling the unobserved pixel-level population counts and, then the conditional distribution of land use given the associated count. The model for the latent population counts leads to a model for the observed counts. With two sets of spatial effects, one associated with the town population counts, the other with pixel level land-use model, a multilevel hierarchical model results.

We show how such models can be fitted using Gibbs sampling (Gelfand and Smith, 1990), thus allowing full inference of all of the modeling levels (cf. Gelfand et al., 2005). The misalignment problem extends recent work of Mugglin and Carlin (1998), Mugglin et al. (2000) and Gelfand et al. (2003). In the present context the other explanatory variables we employ are also measured at the (l km × 1 km) pixel-level. However, if this were not the case, a general strategy is as follows. Identify the data layer at the finest spatial resolution. For each remaining layer create an individual overlay on this one. For each overlay, carry out a rasterization so that each areal unit of the latter is contained on one and only one unit of the former. Then model the joint distribution of all the variable layers at this highest resolution. Such an approach yields a model for each data layer at its observed scale. Note that it is not necessary to align all pairs of data layers. Rather, we need only align each layer with the one at the finest spatial resolution.

Two related implementation issues arise under our modeling approach. First, for the response data we work with, due to the presence of cloud cover, land-use classification is missing for approximately 8% of the roughly 47,000 pixels. Thus, we introduce an imputation approach, i.e., a spatial multiple imputation (Schafer, 1997), to handle this missing data. Second, our modeling requires commitment of considerable computational effort (tailored programs for the individual application) and computer run time. To justify this commitment we offer model comparison revealing the inferential benefits of our modeling relative to simpler model specifications, which can be fitted more easily and quickly.

4. The dataset

The data used in developing the deforestation model are from Madagascar, an area of the world designated as particularly high priority for conservation efforts. It is recognized as 1 of 25 mega-diversity countries in the world (Meyers et al., 2000) and the eastern tropical wet forest in Madagascar is globally one of 12 rain forest “hot spots”. At the same time, the forests in Madagascar have been characterized as under tremendous threat from deforestation (USDA Forest Service, 2000; Sussman et al., 1994; Oxby, 1985). As a consequence many of the most well-known plant and animal species are listed by the International Union for the Conservation of Nature (IUCN) as globally threatened.

There is considerable difference of opinion as to how much of the original forest in Madagascar is left as we pointed out earlier, but estimates range down to as little as 10%, and perhaps only 25% of what is left is considered primary or undisturbed forest (Sussman et al., 1994; USDA Forest Service, 2000). Of these systems, the coastal forests are the most threatened. Estimated deforestation rates for Madagascar are quite variable: less than 1% per year to 5% or more. Some have estimated that within 20 years, little or no forest will remain outside protected areas (currently about 1.85% of the total land surface). But, as mentioned earlier, most of these estimates of forest cover and deforestation rates are speculative and controversial. Concerns over deforestation and attempts to estimate rates is nothing new. In 1921 Perrier de la Bâthie estimated that between 1900 and 1920, 49,000 km^2 of forest had been lost. Early in the 19th century, the first King of Madagascar proscribed the clearing and burning of forest (Verin and Griveaud, 1968).

The focal area for this study is the wet tropical forest biome within Toamasina (or Tamatave) Province, Madagascar. This province is located along the East Coast of Madagascar, and encompasses the greatest extent of tropical rain forest on the island nation (cf. Fig. 1). The aerial extent of Toamasina Province is roughly 75,000 km^2. To model forest cover for the province, we obtained or constructed five digital geospatial layers representing: town boundaries (with asociated 1993 population census data), elevation, slope, road and transportation networks (1993) and land cover.

4.1. Forest classification

A single NOAA-11 AVHRR, 10-day composite image for September 21, 1993 was obtained from the USGS EROS Data Center (http://edcwww.cr.usgs.gov). Three reflective bands (1, 2 and 3) plus a Normalized Difference Vegetation Index (NDVI) (bands $3-2\backslash 3+2$), were presented to an unsupervised Iterative Self Organizing (ISODATA) clustering algorithm, within ERDAS Imagine v.8.4 software. This clustering output 30 initial classes, which were interpreted using expert personal knowledge of the cover types, their composition, distribution and location within the region. Two Landsat 5 Thematic Mapper TM images (September 15, 1993 and November 21, 1994), a set of aerial photographs from the 1940s and 1960s together with extensive independently collected ground-truthing data were used to aid fine scale pattern delineation. The 30 classes were iteratively assessed, reclassified with a supervised, maximum-likelihood process and reduced to the final 5 used for the modeling: mature forest, secondary patch forest, cleared forest, non-forest and missing values (for small patch areas obscured by clouds or cloud shadows, haze, etc.). We believe this provides the most accurate characterization of the landscape at this scale available. The vector maps of forest cover currently available (Faramalala, 1995; Du Puy and Moat, 1999) are overly smoothed and in many cases report forested and non-forested areas that based on our ground truth information, are misclassified.

4.2. DEM generation

The quality of existing global topographic information (Gtopo30, Digital Chart of the World) we found to be inadequate, containing numerous artifacts and noise. Instead we use vector elevation contour data (100 m) digitized by the Madagascar geodetic survey office (Intitut National de Geodesie et Cartographie, or Foiben Taosarin-tanin 'I Madagasikara (FTM)). These were used to generate a continuous, 1 km × 1 km grid surface, utilizing a combination of Surfer v. 6.04 (Golden Software, Golden CO) and Arc View Spatial Analyst v. 2.0 (ESRI, Redlands CA) software. Spatial misalignments and misregistration errors were corrected to produce the final 1 km^2 elevation grid layer. In turn, this layer was used to calculate a slope (degrees) layer, using Arc View Spatial Analyst v. 2.0.

4.3. Town, population and other cartographic data

Township (i.e., "firaisana") boundary information was digitized for all 162 firaisana in Taomasina province by the Madagascar geodetic survey office (FTM). Census data were obtained from the United Nations via the Madagascar census bureau (Direction et Statistique Sociales) for each firaisana for 1993. Vector point data for various population center features (i.e., village/hamlet localities or "fokontany"), firaisana "headquarter" villages, fivindronana headquarter towns, and the provincial capitol) were digitized and supplied by FTM. Road and tracks were also digitized and supplied by FTM (primary routes (sealed roads) and railroads, motorable (in dry season) tracks, rough ox cart tracks, and primary footpaths). These vector layers were all rasterized to 1 km^2 grids. Grid cells fell along town boundaries were assigned to a particular town by majority area rule. Grid cells with one type of road were assigned the appropriate value. When two or more road/path network types fell in a particular grid cell, the more developed road value was assigned to the pixel.

The final layers were rasterized to 1 km and clipped to the extent of the Town boundaries in the province. All islands were clipped as was the one town that covered less than 1 km^2. All data were obtained in, or projected to, Laborde (m), WGS84 coordinate values.

Following clipping, 159 townships and 74,607 1 km × 1 km pixels remained in our data set. Fig. 2 shows the town level map for the towns in the study region. In the western portion of the province an escarpment demarcates the eastern wet forest or potential forest biome from the western, seasonally dry grassland/savannas mosaic. Visual inspection of the data layers prior to analyses revealed some differences between the northern and southern parts of the province. The North had fewer population centers most of which tend to be clustered near the coast, and larger forest patches, while the South had more populations centers scattered across the landscape, with many smaller forest patches and more extensive road development, including primary routes to the national capital, west of the study region. With concern that pattern and process might differ between these regions and with interest in making comparisons between them, we created two disjoint regions. In particular, we excluded the western (non-forested majority) towns and introduced, between North and South, a buffer region at least two towns wide (to provide separation of North and South spatial effects). This yielded the North and South regions identified in Fig. 2. The North included 46 towns encompassing 22,347 km^2 pixels and total population 707,786; in the South there are 66 towns encompassing 24,623 such pixels and total population 664,066.

We note that in creating these two disjoint subregions, the concern was not to strive for homogeneity of spatial pattern of the sub-regions. Rather, due to the fairly simple mean specification of the model described below, we were concerned that omitted or unobserved variables which carry spatial information may differentially affect land use in the sub-regions. This implies that the spatial model for random effects (which are introduced as surrogates for these variables) may differ among sub-regions. Moreover, the regression coefficients of the observed explanatory variables may also vary among sub-regions. In general, the message is that if, a priori, one suspects pattern or process is operating differentially in portions of the study

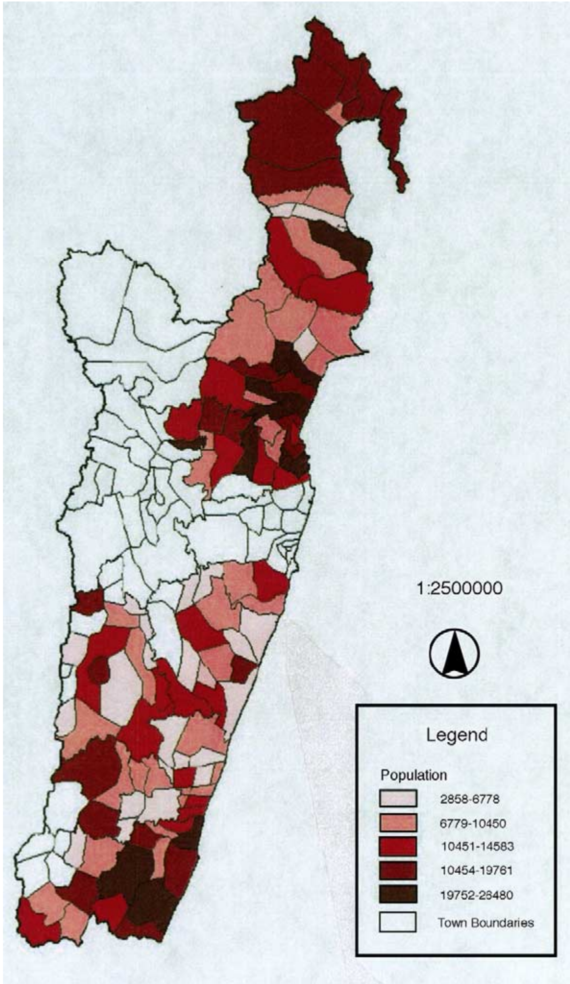

Fig. 2. Northern and southern regions defined within the study region. All town boundaries are shown. Population size categories are shown for northern and southern towns studied.

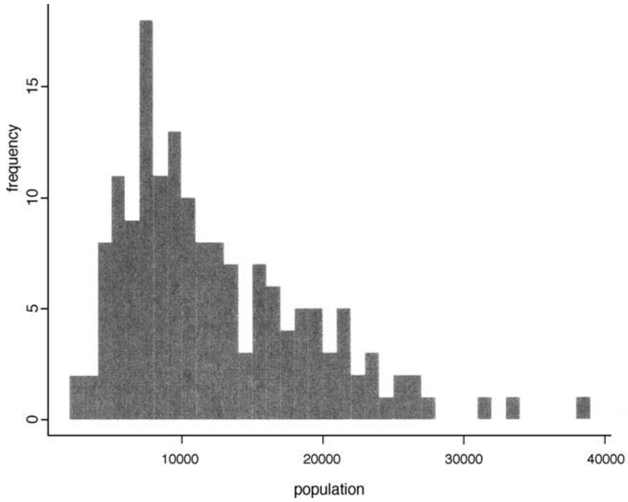

Fig. 3. Histogram for population for the 159 towns in the full study area.

region, it is prudent to fit individual models to these portions.

The GIS data layers used in the modeling are shown in Figs. 2–7. In Fig. 2, town population data for North versus South is overlaid on the town boundaries for the entire province. Fig. 3 provides a histogram of the population data for all towns included in the model.

The ordinal land-use classifications for the northern and southern region are shown in Fig. 4: the most degraded land is shown as non-forest (class 1), followed by cleared and degraded (potential) forest lands (class 2), secondary patch forest (class 3) and mature forest (class 4) and the missing value class. Relative contribution (in percentage) of each class is provided in Table 1. Fig. 5 provides grey scale maps for elevation while Fig. 6 provides grey scale maps for slope.

Road classification was modified prior to full model development. Initially, we had five road types assigned to pixels, viz. no roads, footpath, rough track, motorable track and primary route or rail road (Fig. 7). However, due to the sparseness of counts in some of the categories, we needed to reduce this to a binary road classifications. In fact, because the sparseness patterns differed between the North and the South, we used differently defined binary classifications for the two regions. In the North, no roads was one of the binary road classifications while the other four road types formed the second classification. In the South, no roads and footpath formed one category while the other three road types were lumped into a second category.

5. Model details

We model the joint distribution of land use (L) and population count (P) at the pixel level (1 km^2). Let L_{ij} denote the land-use value for the jth pixel in the ith township and let P_{ij} denote the population count for the jth pixel in the ith township. The L_{ij} are observed but only $P_i = \sum_j P_{ij}$ are observed at the town level. We collect the L_{ij} and P_{ij} into town level vectors $\boldsymbol{L}_i$ and $\boldsymbol{P}_i$ and overall vectors $\boldsymbol{L}$ and $\boldsymbol{P}$. We also observe at each pixel elevation, E_{ij}, slope, S_{ij} and road classification R_{ij}. Lastly, let δ_i denote a spatial random effect for township i. We elaborate these effects below.

Table 1
Contribution of land-use classes in percentage

	Class 1 (non-forest)	Class 2 (cleared/ degraded forest)	Class 3 (secondary patch forest)	Class 4 (mature forest)	Missing
North	2.5	30.7	17.2	43.5	6.1
South	10.7	42.5	11.1	26.2	9.5

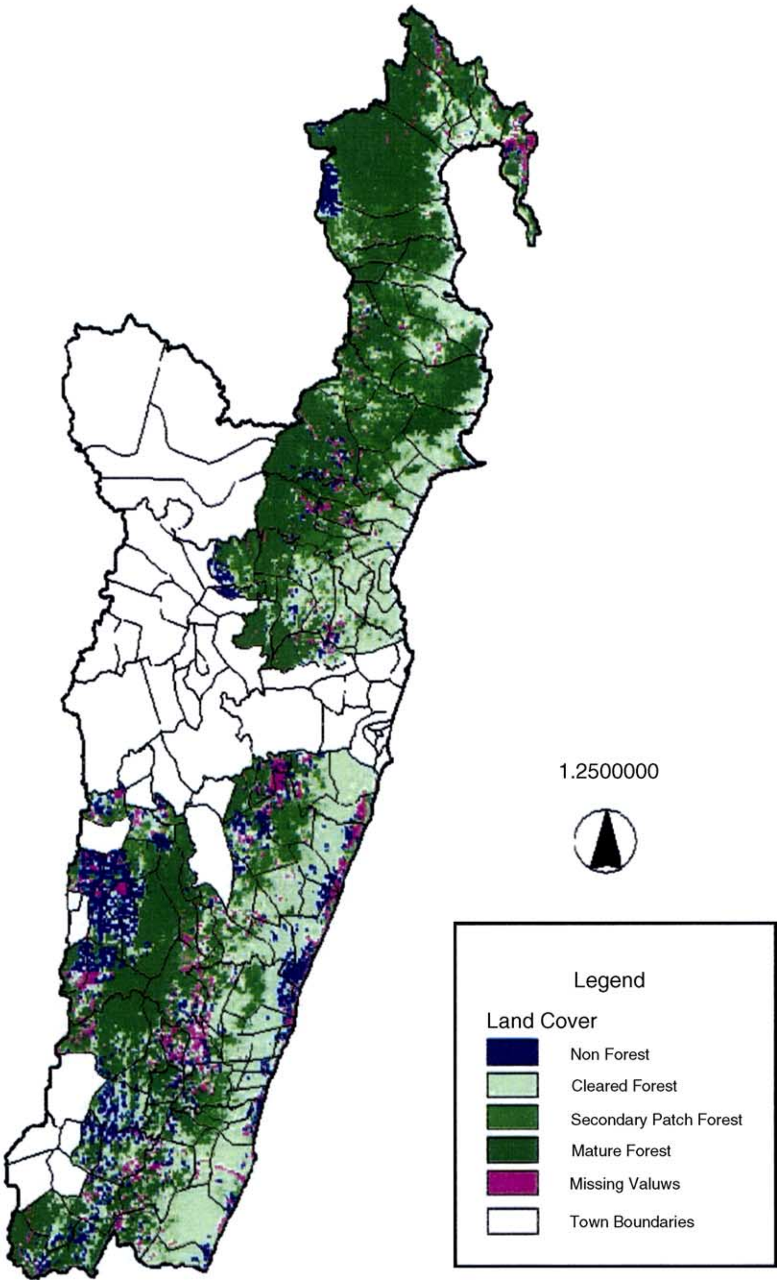

Fig. 4. Ordinal land-use classifications for the study regions.

In specifying the joint distribution, $f(\boldsymbol{L}, \boldsymbol{P}|\{E_{ij}\}, \{S_{ij}\}, \{R_{ij}\}, \{\delta_i\})$, we factor this joint distribution as

$$f(\boldsymbol{P}|\{E_{ij}\}, \{S_{ij}\}, \{R_{ij}\}, \{\delta_i\})f(\boldsymbol{L}|\boldsymbol{P}, \{E_{ij}\}, \{S_{ij}\}, \{R_{ij}\}). \quad (1)$$

We condition in this fashion since one of our objectives is to see if population provides a significant effect for land use. Such explanation is evidently not causal; we could equally well-condition $\boldsymbol{P}$ on $\boldsymbol{L}$.

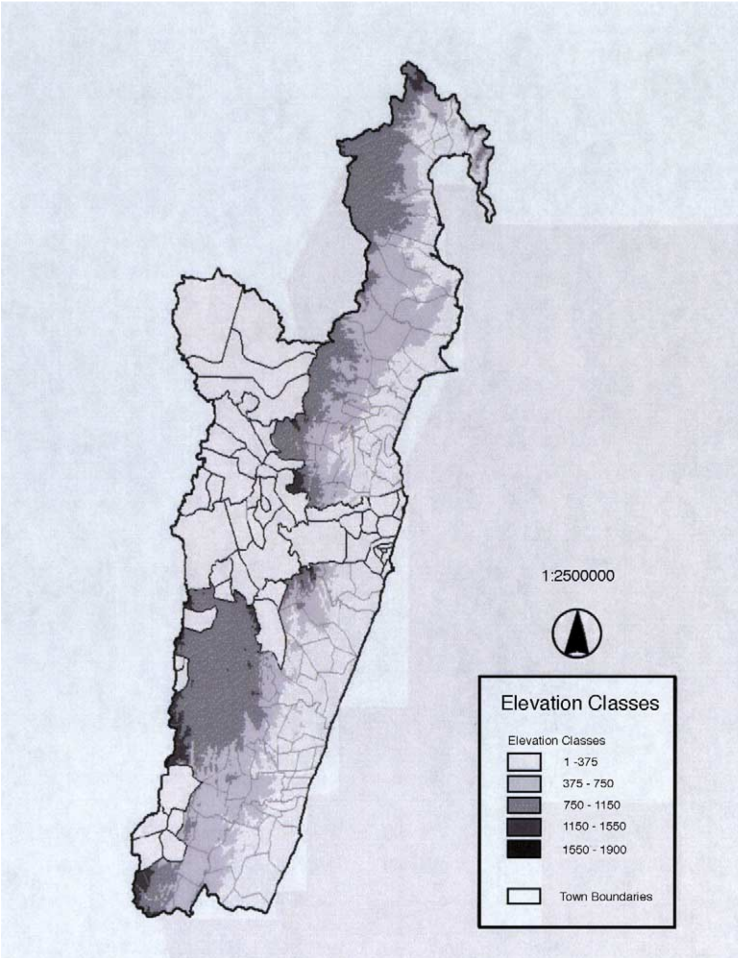

Fig. 5. Grey scale elevation maps for the study regions displayed in a limited number of classes here.

Turning to the first distribution in (1), the population model, we assume that the P_{ij} are conditionally independent given the E's, S's, R's and δ's, i.e., we write

$$f(\boldsymbol{P}|\{E_{ij}\},\{S_{ij}\},\{R_{ij}\},\{\delta_i\}) = \prod_i \prod_j f(P_{ij}|E_{ij}, S_{ij}, R_{ij}, \delta_i). \quad (2)$$

Moreover, since P_{ij} is a population count and since many of these (unobserved) counts will be sparse for this region at this scale, we assume P_{ij}: Poisson (λ_{ij}) where

$$\log \lambda_{ij} = \beta_0 + \beta_1 E_{ij} + \beta_2 S_{ij} + \beta_3 R_{ij} + \delta_i \quad (3)$$

Since the P_{ij}'s are conditionally independent Poisson variables, summing over j yields P_i: Poisson (λ_i) where $\log \lambda_i = \log \sum_j \lambda_{ij} = \log \sum_j \exp(\beta_0 + \beta_1 E_{ij} + \beta_2 S_{ij} + \beta_3 R_{ij} + \delta_i)$.

In (3), the β''s are regression coefficients while, again, the δ_i's are township level spatial random effects. They are intended to capture anticipated spatial similarity for neighboring townships with regard to population. Unlike usual random effects which are assumed to be independent and identically distributed normal random variables, the distributions of the δ_i's are specified conditionally given all of their neighbors. That is,

$$\delta_i|\delta_j,\ j \neq i \sim N\left(\frac{\sum_j \omega_{ij}\delta_j}{\sum_j \omega_{ij}}, \frac{\tau^2}{\sum_j \omega_{ij}}\right)$$

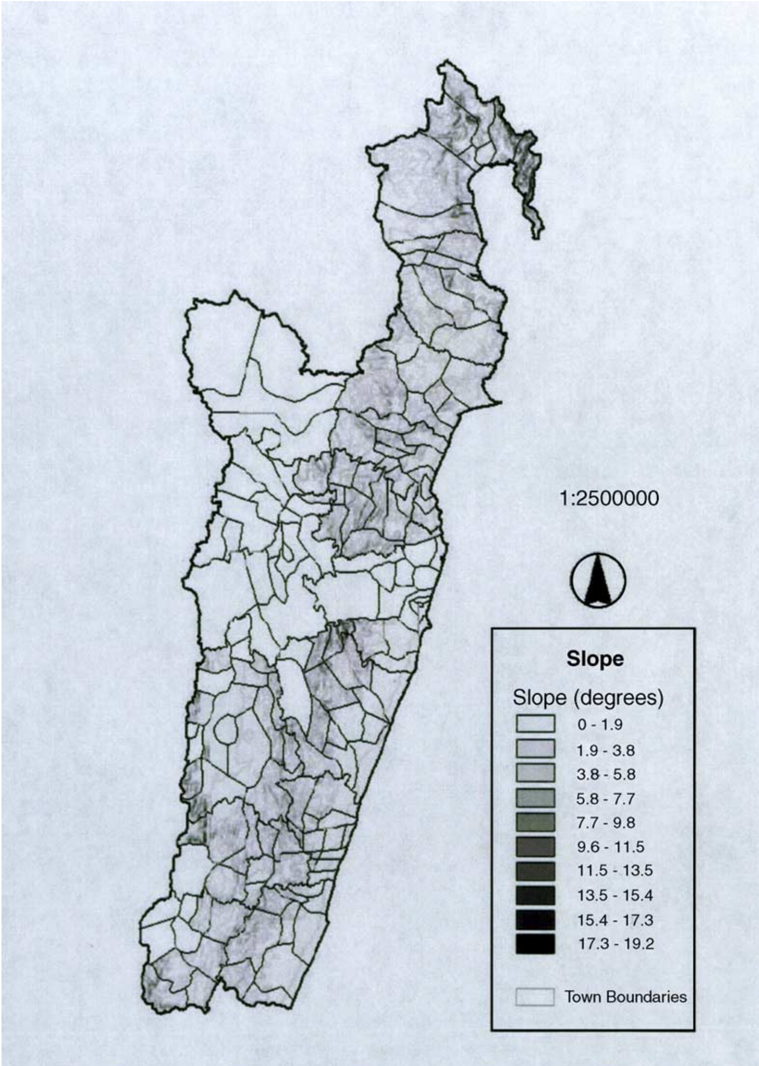

Fig. 6. Grey scale slope maps for the study regions.

where $\omega_{ij} = 1$ or 0 according to whether or not town j is contiguous with town i and τ^2 is the spatial variance component. In other words, δ_i is expected to vary about the average of its neighbors. Such models are referred to as conditionally autoregressive (CAR) models. See, for example, Besag (1974) or Cressie (1993) for further details. We note that since the P_{ij} are not observed we cannot introduce pixel level spatial effects into the population model. We cannot adjust the mean of population counts we have not seen; unstable model fitting results. So, spatial adjustment to λ_{ij} will be identical for all pixels within a town and will be similar for pixels arising from adjacent towns. Also, since the P_{ij} are conditionally independent random variables with $\sum_j P_{ij} = P_i$, we have $\{P_{ij}\}|P_i$: multinomial(P_i; $\{\gamma_{ij}\}$) where $\gamma_{ij} = \lambda_{ij}/\lambda_i$.

In the second term in (1), the land-use classification model, we assume conditional independence of the L_{ij} given the P's, E's, S's and R's. To handle the ordinal nature of the L_{ij}'s, we follow Albert and Chib (1993), introducing a latent variable W_{ij} associated with each L_{ij}. W_{ij} is conceived as a continuous random variable

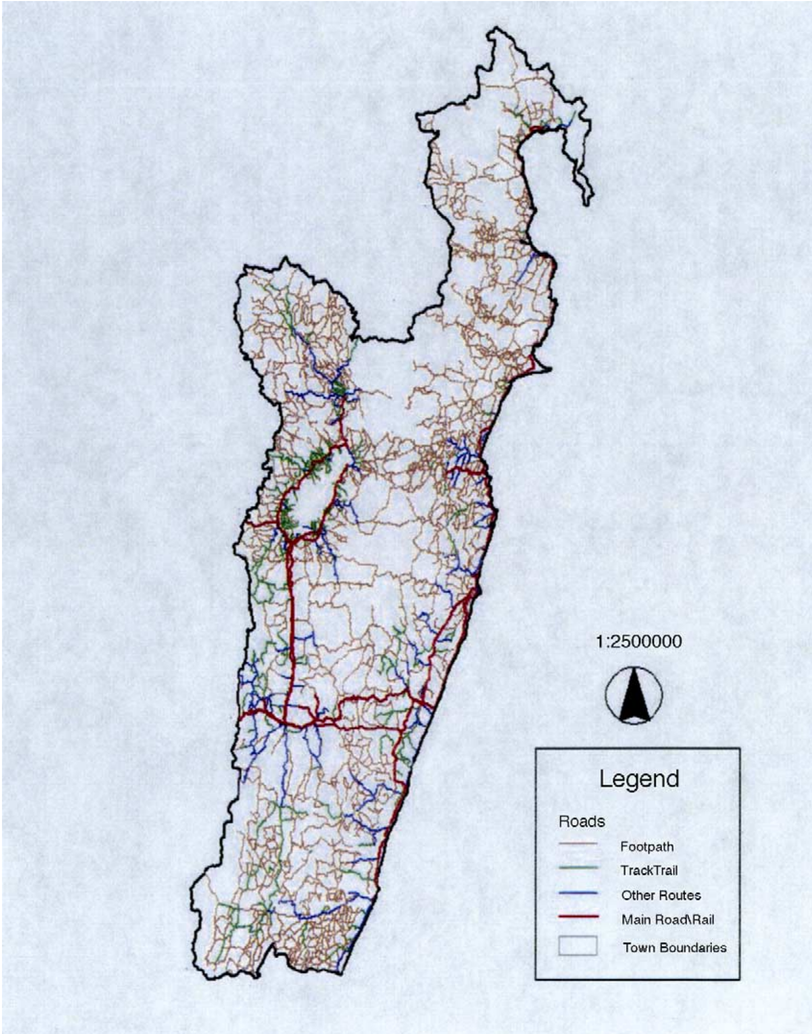

Fig. 7. Vector road classification for the entire province.

on the real line, interpreted as the extent of forest cover. We then imagine that the four ordinal land-use classifications cut the real line into four intervals so that $L_{ij} = l$ if $W_{ij} \in (\gamma_{l-1}, \gamma_l)$, $l = 1, \ldots, 4$ with regard to degradation where $\gamma_0 = -\infty$, $\gamma_4 = \infty$. Hence, $\pi_{ijl} \equiv P(L_{ij} = l) = P(W_{ij} \in (\gamma_{l-1}, \gamma_l))$, $l = 1, \ldots, 4$, provides the four land-use classification probabilities at each pixel in each town. In other words, L_{ij} is a multinomial trial with possible outcomes $l = 1, \ldots, 4$ having respective probabilities π_{ijl}, $l = 1, \ldots, 4$. The W_{ij}s can only be identified up to translations. That is, the W_{ij}'s are located *relative*to each other on this imaginary scale but the data cannot *absolutely* place the W_{ij}'s. To do so, without loss of generality, we can set the cut point $\gamma_1 = 0$ but γ_2 and γ_3 are unknown. Hence, given W_{ij} and the γ's, L_{ij} is determined, i.e., the conditional distribution degenerates to a particular l given W_{ij}. Conversely, given L_{ij} and the γ's, W_{ij} is restricted to an interval. Thus, the land-use classification model becomes $f(\boldsymbol{L}, \boldsymbol{W}|\{E_{ij}\}, \{S_{ij}\}, \{R_{ij}\}, \{P_{ij}\}, \gamma_2, \gamma_3)$ which we can write as

$$f(\boldsymbol{L}|\boldsymbol{W}, \gamma_2, \gamma_3) \cdot f(\boldsymbol{W}|\{E_{ij}\}, \{S_{ij}\}, \{R_{ij}\}, \{P_{ij}\}). \quad (4)$$

Again, due to the conditional independence, the first distribution in (4) can be written as $\Pi_i \Pi_j f(L_{ij}/W_{ij}|E_{ij}, \gamma_2, \gamma_3)$. The second becomes $\Pi_i \Pi_j f(W_{ij}|E_{ij}, S_{ij}, R_{ij}, P_{ij})$ where we use a linear regression model,

$$E(W_{ij}) = \alpha_0 + \alpha_1 E_{ij} + \alpha_2 S_{ij} + \alpha_3 R_{ij} + \alpha_4 P_{ij}. \quad (5)$$

However, note that because the W_{ij}'s are not observed, their latent scale is not identifiable. We could multiply each by a constant, multiply the α's by this constant and multiply the standard deviation of W_{ij} by this constant and the modeling would be unaffected. Hence, without loss of generality, we set $\mathrm{var}(W_{ij}) = 1$. Note further that we cannot introduce spatial random effects (in fact any random effects at the pixel level) into the land-use model. Were we to propose including effects say φ_{ij} in (5), since P_{ij} is not observed, the data would not be able to separate P_{ij} and φ_{ij} in the component $\alpha_4 P_{ij} + \varphi_{ij}$. Nonetheless, the δ_i's in (3) induce spatial smoothing in explaining the P_{ij}'s which, in turn, produces some spatial smoothness in explaining the W_{ij}, hence L_{ij}. A simpler version of the model in (1)–(5) is presented with some analysis along with a more technical perspective in Agarwal et al. (2002).

We conclude with several remarks:

Remark 1. A more customary modeling approach to explain the L_{ij} would be a non-hierarchical categorical regression. For instance, we could model $\mathrm{logit}\frac{\pi_{ij,l-1}}{\pi_{ij,l}}$, $l = 2, 3, 4$ as a linear model in E_{ij}, S_{ij} and R_{ij}. Such models can be routinely fitted using standard statistical software packages but since we do not have P_{ij}, what should we do? Set $P_{ij} = P_i$? Set $P_{ij} = P_i$ (area of ijth pixel)/(area of township i)? These spcifications are equivalent since all pixels are the same size. However, the assumption of a uniform distribution across townships is inappropriate as Fig. 9 below reveals. By introducing the latent P_{ij}'s, the E_{ij}, S_{ij}, R_{ij}, and δ_i allow us to learn about them while reflecting our uncertainty in their actual values.

Remark 2. More generally, note how our modeling explicitly handles the misalignment issue. To build a model at the town level for the P_i, what should we use for E_i, S_i, and R_i? To build a model for L_{ij}, we should introduce a P_{ij}.

Remark 3. It is perfectly fine to use the covariates E_{ij}, S_{ij}, and R_{ij}, to characterize the joint distribution of L_{ij} and P_{ij}. However, with regard to conditional and marginal specifications, if we use these covariates to explain P_{ij}, should we also use them to explain L_{ij} given P_{ij} (or perhaps, vice versa)? In familiar normal linear modeling, we would introduce redundant parameters in the marginal model for L_{ij}. The marginal model is over-parametrized, the regression coefficients are not identified. Within the Bayesian framework, no problem arises. We would find little learning about the components of the coefficient vector but no problem with the overall coefficient. In the present situation, with a Poisson model for the P_{ij}'s, the identifiability problem does not arise; in the marginal model for L_{ij}, we have both an additive form and a multiplicative form in the covariates. Perhaps more importantly, the marginal model for L_{ij} would not involve P_{ij}. We are explicitly interested in a conditional explanation for L_{ij} given P_{ij} and the other covariates.

6. Fitting and inference for the model

The model defined by (1)–(5) is referred to as a multilevel or hierarchical specification. In fact, at the lowest level we have the observed land-use classifications. At the level above we have the latent W's. At the top level we have the latent pixel level populations constrained by the observed township populations. This model is of high dimension with unknowns, $\{W_{ij}\}$, $\{P_{ij}\}$, $\{\delta_i\}$, α, β, γ_2, γ_3 and τ^2. Fitting of this entire model is feasible only within a Bayesian framework. That is, we have already provided the joint distribution for L, W and P given the model parameters. If we view these latter unknowns as random and add a so-called prior distribution for them we have a complete model specification, i.e., a specification for the joint distribution of all the variables in the hierarchical model. With this joint distribution we can provide inference regarding any aspect of the model. The prior distribution for the δ's was described above. For α, β, γ_2, γ_3 and τ^2, the general approach is to use whatever prior or partially data-based information we may have regarding these unknowns to obtain some idea of what range they are likely to fall in. We use this range to develop proper prior distributions, which are as non-informative (i.e., as vague) as possible (in order to let the observed data drive the inference) while still retaining stable computation. Details of the prior specifications are supplied in Appendix A.

The resulting model is fitted using simulation-based methods, i.e., Gibbs sampling (Gelfand and Smith, 1990) and Markov chain Monte-Carlo methods (see, for example, Gilks et al., 1996). The output from such simulation is sampled from the joint posterior distribution of all model unknowns. Unfortunately, such model fitting requires considerable effort and time. (In fact, we summarize the required full conditional distributions and provide some pseudo-code for implementing the sampling in Appendix B.) The reward is exact inference (without relying on possibly inappropriate asymptotics) and more accurate measurement of variability by capturing the uncertainty regarding the model unknowns, which is not obtained using classical statistical approaches. In fact, we know of no other way to fit models (1)–(5). Simplified versions may be fitted using the E–M algorithm (Dempster et al., 1977) to obtain likelihood-based inference. But then we have concerns regarding associated asymptotic variance estimates. It may also be possible to fit the model stages separately rather than simultaneously. Again, we are concerned that this will misrepresent the extent of variability. These relate to issues of model choice, which we discuss below.

7. Imputing missing land-use classification

Table 1 shows the distribution of land-use classifications for the North and South regions. Note that 6.1% of classifications are missing (primarily due to cloud cover) in the North and 9.5% are in the South. One of our fundamental modeling assumptions is that each pixel belongs to one and only one town and the pixel level population, when aggregated over a town, gives the known town level population. Thus, we cannot delete pixels with missing land-use classification from our analysis. Rather than relying on subjective or ad hoc determination of these missing pixel values, we elect to impute missing land-use values in order to have a complete set for fitting our hierarchical model. In fact, the imputation provides an entire set of missing land-use classifications using a neighbor-based joint spatial distribution for the pixel values. Such a distribution should support (but not require) land use for a given pixel to be similar to that of its neighbors. Moreover, we do several such imputations to see the sensitivity of our resulting inference to the imputation, to better assess the variability in the model unknowns. Various imputation models could be adopted; we employed a convenient Potts model (Potts, 1952; Green and Richardson, 2002). Details are supplied in Appendix C. We used three imputations, finding negligible differences in the resulting parameter estimates across these imputations.

8. Model determination

The model formalized in (1)–(5) is elaborate and challenging to fit. Is the effort justified? Here we consider three simplifications resulting in easier to fit models and, using a model choice criterion, show that they are not as good.

For instance, a crude model could ignore the population model (3). Instead, we could fit the land-use model in (5) inserting for the unknown P_{ij}'s an areally allocated portion of the total township population to each pixel within the township. In other words, if there are n_i pixels in township i, we set $P_{ij} = P_i/n_i$. Hence we have only the one stage model in (4) and (5) to fit.

An improvement in this naive approach first fits the model in (3) (with or without spatial effects). This fitting considers the P_i's as the only data, the P_{ij}'s are latent variables, and estimates the P_{ij}'s along with β (and δ if spatial effects are included). The resultant fitting provides posterior means $E(P_{ij}|P_i)$ which could then be inserted into the land-use model (5). These expected values are anticipated to provide better estimates of the true P_{ij} than those from the crude allocation. Furthermore, with the inclusion of the δ_i, the $E(P_{ij}|\{P_i\})$ will tend to be smoother than without. But regardless, we are still failing to capture the uncertainty in not knowing P_{ij}; we will underestimate the variability in the overall model. We note that in the absence of spatial effects, our fitting approach is analogous to the use of the E–M algorithm to obtain maximum likelihood estimates (Dempster et al., 1977) with incomplete data (the P_{ij} are missing). In fact, we can also obtain the customary likelihood inference using the E–M algorithm. Again, concern is with the resultant interval estimates, i.e., the approximate normality assumption on which they are based and the goodness of the variability estimates (estimated standard errors) which they require.

To compare the foregoing models we use a version of the posterior predictive loss approach in Gelfand and Ghosh (1998). Using squared error loss for con-

Table 2
Model comparison for the North and South regions

Model	North			South		
	G	*P*	*D*	*G*	*P*	*D*
I (without population)	10949.68	11911.42	22861.10	15317.60	16788.42	32106.02
II (without spatial)	10776.78	11489.25	22266.03	14667.41	15407.17	30074.58
III (full bayesian)	10751.74	11460.42	22212.16	14670.59	15404.39	30074.98

G_i: goodness of fit term; P_i: penalty term; D_i: combined.

venience this approach provides a criterion, which is composed of two parts. One measures goodness of fit (G) of the model; the other is a penalty term (P), penalizing model complexity. In other words, more complex models will tend to fit better but should be penalized for their "size" in order to encourage parsimony.

Specifically,

$$G = \sum_{i=1}^{159} \sum_{j=1}^{n_i} \sum_{l=1}^{4} (L_{ijl,\text{obs}} - \pi_{ijl})^2$$

and

$$P = \sum_{i=1}^{159} \sum_{j=1}^{n_i} \sum_{l=1}^{4} V(L_{ijl})$$

where $L_{ijl,\text{obs}} = 1$ if $L_{ij,\text{obs}} = l$, 0 otherwise, $\pi_{ijl} = P(L_{ij} = l|L_{obs})$ and $V(L_{ijl})$ is the predictive variability of $L_{ij\ell}$, i.e., $\pi_{ijl}(1 - \pi_{ijl})$. Recall that for each i, j, one $L_{ijl,\text{obs}} = 1$ and the remaining three are 0. Under G, a model is incrementally rewarded at an i, j pair if its π_{ijl}'s (which also sum to 1) behave similarly to the $L_{ijl,\text{obs}}$. Under P, a model is rewarded at i, j if it produces "informative," i.e., extreme π_{ijl}'s . In calculating G and P we could sum over only the observed L_{ij}'s or also over the imputed L_{ij}'s. Since the imputation was implemented apart from the model fitting, we chose the latter.

It is useful to record both G and P (in fact, in some cases it might be useful to plot the components of G and P. However, if we seek to reduce model choice to a single number, then we choose the model which minimizes $D = G + P$. Table 2 presents the values of D, G and P for both the North region and the South region for Model I (simple allocation, without population model) Model II (improved model, includes population model, but not spatial effects) and Model III (the full Bayesian model in (1)–(5)). Clearly, Models II and III are preferred to Model I and, at least in the North, Model III is preferred to Model II.

9. Model results

Table 3 provides point estimates and 95% credible intervals for both the population and land-use models. For comparison, Table 4 presents point estimates and associated interval estimates for Model II fitted using an E–M algorithm. We see that with regard to the population model, the two models provide very similar point estimates but that the approximate likelihood in-

Table 3
Median and 95% credible interval (in parentheses) for the full Bayesian Model III

	North	South
Population model parameters		
β_1 (elevation) $\times 10^{-4}$	−8.210 (−8.412, −8.038)	−6.487 (−7.028, −5.986)
β_2 (slope)	.0704 (.0682, .0725)	−.0226 (−.0302, −.0150)
β_3 (roads)	1.157 (1.151, 1.163)	.2124 (.1922, .2314)
τ_{δ^2}	1.503 (1.058, 2.423)	1.791 (1.297, 2.455)
Land-use model parameters		
α_1 (elevation) $\times 10^{-3}$	4.144 (4.024, 4.242)	1.942 (1.874, 2.018)
α_2 (slope)	−.0089 (−.0198, .0020)	.0415 (.0306, .0526)
α_3 (roads)	−.5931 (−.6649, −.5232)	−.4587 (−.5373, −.3892)
α_4 (population)	−.0101 (−.0119,−.0081)	−.0028 (−.0033, −.0024)

Table 4
Point estimates and 95% interval estimates (in parentheses) obtained from the EM algorithm (Model II)

Region	North	South
Population model parameters		
β_1 (elevation) $\times 10^{-4}$	−8.217 (−8.287, −8.147)	−6.397 (−6.467, −6.327)
β_2 (slope)	.0703 (.0695, −0711)	−.0241 (−.0253, −.0229)
β_3 (roads)	1.157 (1.154, 1.160)	.2182 (.2151, .2213)
Land-use model parameters		
α_1 (elevation) $\times 10^{-3}$	4.304 (4.292, 4.316)	2.084 (2.076, 2.092)
α_2 (slope)	−.0075 (−.0191, .0041)	.0621 (.0499, −0742)
α_3 (roads)	−.7377 (−.8232, −.6522)	−.4744 (−.5548, −.3940)
α_4 (population)	−.0019 (−.0033, −.0005)	−.0024 (−.0028, −.0020)

ference in Model II produces much tighter confidence intervals revealing the significant underestimation of uncertainty. For the land-use model, at least for elevation, this continues to be the case. Thus, for the remainder of this section we describe the results for Model III.

For this model, lower elevation and presence of road networks are associated with population settlement in the North as well as the South with the effects being more significant in the North. In the South, lower slope is associated with higher expected population, which is consistent with what we anticipate. In the North, however, higher slope is associated with higher expected population, contrary to what one might expect. Turning to the land-use model, the coefficient for population although small is significantly negative for the North and the South with the influence being more pronounced in the North. In fact, the interval estimates for the North and South do not overlap. Increased population pressure is associated with increased chances of deforestation with the effect being less in the South. Elevation is associated with forest cover with the effect being more pronounced in the North. Here too, the interval estimates for the two regions don't overlap. Presence of road networks significantly increase the chance of deforestation in the North as well as South. Slope is mixed with regard to significance. While it emerges as insignificant in the North, it significantly increases the chances of forest cover in the South.

Recall that the spatial random effects, the δ_i's, were introduced to provide spatial smoothing to the population model which in turn induces such smoothing to the land-use model. A priori, under the CAR model, all δ_i's have essentially the same distribution and mean 0. A posteriori, spatial pattern emerges. To see this, in Fig. 8 we present a map of the posterior means of the δ_i's.The figure clearly reveals that spatial proximity encourages similarity in the expected spatial effects of population processes. Here, lighter colors indicate a negative expected spatial effect, i.e., the adjustment to the mean population for pixels in that township will tend to diminish population, vice versa for the darker colors.

We have claimed that a model for pixel population such as (3) is more effective than a naive areal allocation of township population to pixels. The results of the model comparison presented in Table 2 certainly support this. Here, we present some further informal support. In Fig. 9 we have created a map of the posterior mean populations (on the square root scale) at the pixel level. The square root scale is employed since this is the customary symmetrizing and variance stabilizing transformation for count data. The smoothing of the population is evident at finer pixel resolutions. Overlaid on this map is the population center dot map. The dots locate the villages, town centers or larger administrative centers. The size of the dot roughly corresponds to population size. Though the dot map could not be reliably used to model pixel level population, it is evident from Fig. 9 that, generally, there is good agreement between the maps.

10. Discussion

Returning to the model issues section, we summarize first what has been accomplished. We have formulated an ordinal land-use classification scale. We have built a phenomenological hierarchical model at the pixel level, which explains land use given population and other covariates as well as population

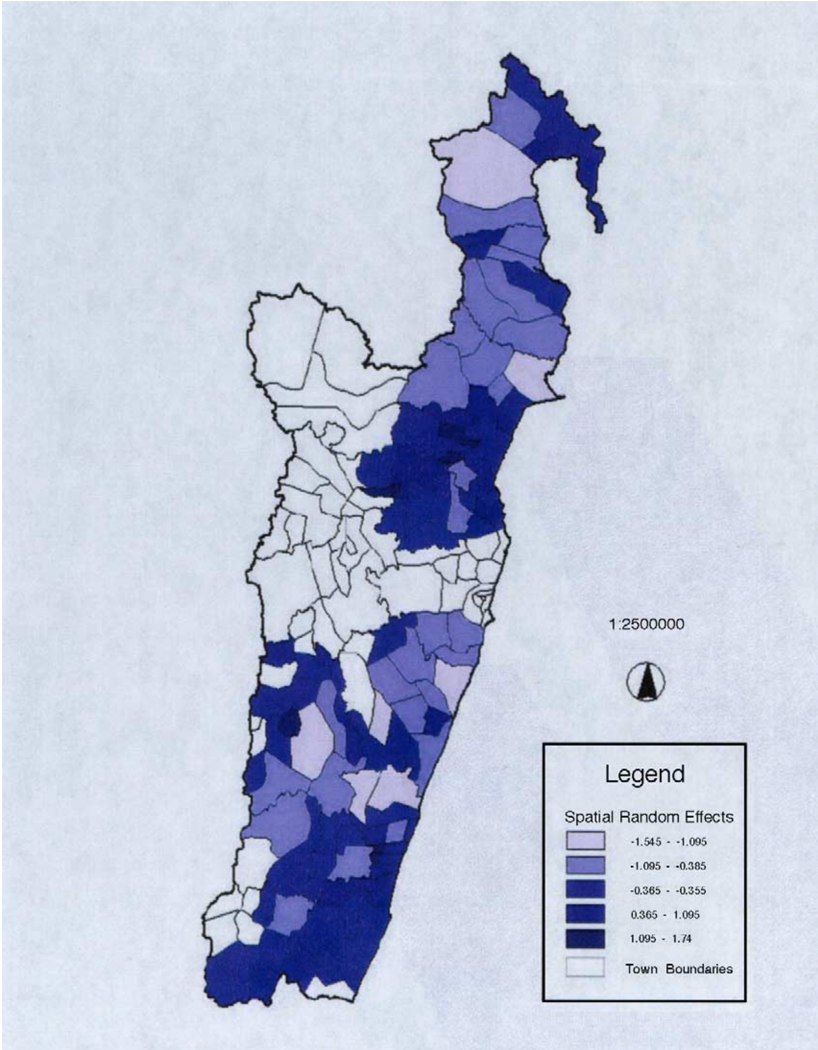

Fig. 8. Population model spatial random effects shown for the study regions.

given other covariates. We have accommodated the misalignment between the population data layer (at the township level) and the land-use data layer (on a 1 km^2 pixel grid). We have employed a spatial imputation (simulation) strategy to handle the missing land-use pixels in an objective fashion. We have introduced spatial smoothing for the population model, the only feasible possibility given the data available. We have uncovered significant explanation in the mean structure from the elevation, slope and roads data layers in both the population and land-use models. In this regard, we have found a significant negative coefficient for population in the land-use model. Increased population pressure increases the chance of forest degradation. We have argued, through a model comparison criterion, that a more sophisticated, hierarchical model, which includes spatial effects explicitly and a population model outperforms models lacking these specifications. We have also argued and demonstrated that our hierarchical model captures uncertainty, which was underestimated by simpler models. Lastly, we have shown the sort of statistical inference that is possible with our hierarchical model, particularly with regard to the spatial smoothing.

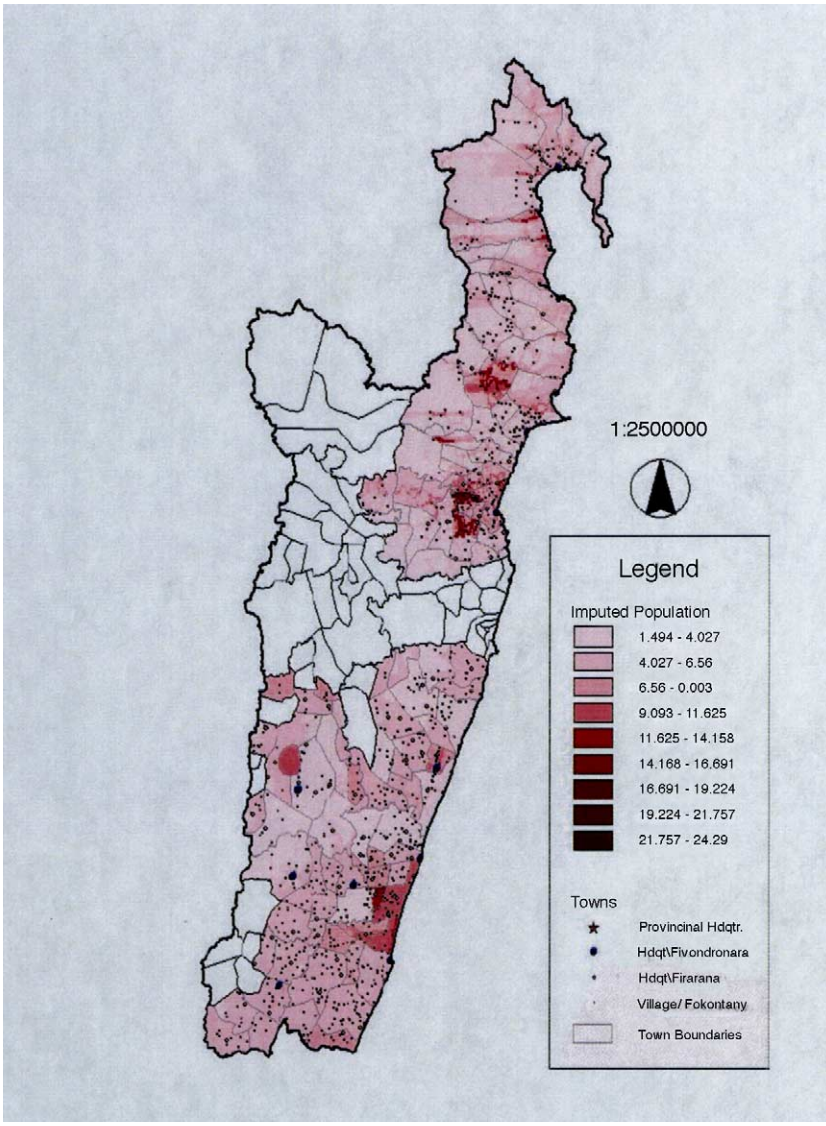

Fig. 9. Expected population for the study regions (on the square root scale) at the pixel level with the population vector point map overlaid.

10.1. Historical context of landuse patterns in Madagascar

Before we discuss the patterns revealed in our analysis, it is instructive to review the historical context of land-use change in eastern Madagascar. Based on our earlier summary of deforestation trends, controversy surrounds issues of rates and timing of forest loss in Madagascar. Claims that most of the forest loss has occurred in recent decades are clearly misleading. The French colonial archives clearly shows concern over forest loss throughout French occupancy (Jarosz, 1993; Perrier de la Bâthie, 1921). In fact the claim, certainly exaggerated, was made that some 70% of the primary forest was destroyed between 1895 and 1925 (Hornac, 1943–1944, cited by Jarosz, 1993). In any case the standard explanations offered for forest loss was the extensive nature of traditional slash-and-burn upland rice culture practiced by the locals, forest concessions encouraged by the French with associated destructive logging practices, and the advocacy of selective forest conversion to plantation and cash crops (Jarosz, 1993; Perrier de la Bâthie, 1921). It is likely that all of these activities would have been encouraged by any site factors that eased access to forested blocks, i.e., adjacency to villages and towns, access by road or even footpath, and lower, shallower slopes. Land tenure systems changed with the 1926 decree that all vacant land be declared state land. Local populations no longer held title to ownership over ancestral lands that might currently be unoccupied (Esoavelomandroso, 1985). Moreover, during the late 19th and early 20th century several laws were passed prohibiting the burning of forests (Jarosz, 1993). Together these laws may well have acted to disenfranchise local populations with respect to traditional land conservation practices, actually exacerbating land degradation.

Yet despite these observations, there are many indications that deforestation processes have been on going for centuries. Few scholars have looked for older records of forest extent in Madagascar. Such exist, in the form of historical accounts of the 17th–19th century (Ellis, 1858; Flacourt, 1658; Grandidier et al., 1920). These narratives make it clear that the distribution of forests were patchy and geographically patterned long before the French colonial period. It is likely that modern patterns of forest distribution reflect, in part, forest distributions of the relatively distant past. The patchy nature of the eastern wet forests was evident to mid-17th century explorers (Grandidier et al., 1920, cf. Flacourt, 1658: Map of Isle Ste. Marie and adjacent Madagascar mainland), as well as the early and mid-19th century travelers (Hébert, 1980; Ellis, 1858). All of these early accounts describe large tracts of deforested landscape from the sea up to 600 m on the mountains. This landscape was clearly a patchwork of cultivated and grazed lands, secondary forests, abandoned lands, and mature forests. Ellis (1858) characterized the limit of closed forest in the mid-1800s as a band along the eastern mountains from 25 to 40 miles wide, not much different from what is found today. There is some evidence that extraction of forest resources was frequent and widespread, together with land clearing along the east coast during the 7th–11th century era (Domenichini-Ramiaramanana, 1988; Dewar, 1997) continuing up to the beginning of the colonial era (Kent, 1992).

The early accounts provide a picture of periodic cycles of landscape clearing for agriculture and abandonment with some evidence for forest regrowth. In part this reflected local cycles of human population growth and decline (Campbell, 1991; Kent, 1992), with the result being a very heterogeneous landscape. The historical perspective is much clearer for the northern section of the region than the southern. This reflects the infrequent visits of Europeans to the southern half of our east coast focal region.

In the central northern area (around Fenerive), there was a shift in human settlements evident in about the 17th century (Wright and Dewar, 2000), with the abandonment of coastal villages, and the establishment of new villages on defensible hilltops at a distance from the coast (cf. Flacourt, 1658: Map of Isle Ste. Marie and adjacent Madagascar mainland). This is also reflected in the narratives of 19th century travelers (cf. Grandidier et al., 1920; Ellis, 1858). Many of these new villages were much larger than previous occupations, so that local population densities were increasing, and may have led to concomitant increase in local anthropogenic effects on the countryside. All of these changes are most plausibly accounted for by an increase in insecurity, and the virtual absence of trading or slaving along the southern coastal area (Kent, 1992), would have meant that population distributions and local environmental effects likely followed divergent paths.

10.2. Forest changes over the past century

Putting the historical information into the context of current land use, it is clear that deforestation has continued up to the present, exacerbated by population growth. The trend has tended to reduce the number and extent of patches of secondary and mature forests in the landscape. This is easily seen on comparing aerial photographs from the 1940s and 1960s, detailed forest patches mapped in the 1960s by Dandoy (1973), and by Benoit de Coignac et al. (1973) about the same time, with recent ground truth and satellite images (data not shown). The result today is a much more homogeneous landscape of degraded lands, with any remaining patches of forest relegated to hill crests and protected valleys. Only along the highest mountain slopes does one find continuous tracts of closed, mature forest. These have retreated in recent decades, but maps of large forest tracts today are surprisingly similar to those of 1969, 1934, 1921 and 1902.

The maps provided by Humbert et al. (1964–1965) have figured prominently in depicting current forest cover for our region (Du Puy and Moat, 1999) and in estimating deforestation rates (Green and Sussman, 1990). Original forest cover was taken from the thumbnail "Types de Végétation" maps, which purport to show potential vegetation in the absence of man (Humbert and Cours Darne, 1965; pp. 152–154). These are conjectured maps with little supporting information, and are primarily constructed from the main "Végétation Naturelle" maps of the east coast. These are used as the baseline "original extent" of forest cover by Green and Sussman (1990), but these provide a misleading benchmark at best. The main "Végétation Naturelle" maps of Humbert and Cours Darne are used to portray current forest cover in the most recent maps of forest cover (Faramalala, 1995; Du Puy and Moat, 1999), as well as forest cover in 1950 (Green and Sussman, 1990). The Humbert and Cours Darne maps were assembled by committee and the forest classification was based on surrogate elevation, bioclimate and edaphic variables with uncertain use of aerial photography, and little or no ground truthing. From our own experiences with aerial photography of the region, vegetation classification without accompanied ground truthing is difficult and misleading. Based on all available information it is apparent that the forest cover maps produced by Humbert and Cours Darne provide a gross over estimate of the extent of forest either currently or for any point over the past half century. Certainly this is true based on even a cursory comparison of the Cartes Forestier with current or recent forest cover on the ground.

10.3. Demographic trends

One can gain some insight on past populations patterns across the landscape from the available demographic maps (de Martonne, 1911; Gourou, 1945; Le Bourdiec et al., 1969). All show a trend North to South that is reflected in the population densities observed in 1993 (cf. Fig. 8). In the South the populations over the past 90 years have tended to be more evenly spread across the landscape. In the North the population centers tend to be more restricted to areas nearer the coast. Over that time, population densities have simply increased across the region.

10.4. Interpretation

How does this information integrate with the results of our spatial analysis of land-use patterns? Population levels are negatively associated with elevation and forested landscape are positively associated with elevation in both the North and the South. This obviously reflects the current and historical location of settlement area at low to moderate elevation and the fact that most of the remaining forest blocks are spatially restricted to the highest mountain slope or locally on hill tops. The effect of elevation is much higher in the North than South. This reflects undoubtedly that fact that large blocks of forest are associated with the higher mountains of this section of the province. Slope is negatively associated with populated blocks in the South but positively associated in the North. This may reflect the fact that settlements were shifted to fortified hill tops (but not at great elevation) in the North beginning in the 17th century (cf. Flacourt, 1658: Map of Isle Ste. Marie and adjacent Madagascar mainland—the center of our North region; no comparable maps are extant for the South). Pixels with hill tops would tend to reflect higher slopes. However, this trend slope was not observed in the South. Rather settlements tend to be spread out more evenly across the landscape over time in areas of convenience, likely avoiding areas of slope and elevation. Slope shows a similar trend in the spatial distribution of forested blocks, North versus South. In the South steep slopes are positively associated with forested blocks. In the North there is no trend with slope. This may reflect the trend of finding some hill tops (and hence steep slopes) occupied by towns and some by forests. An alternative explanation is that the North may represent a rougher topography with greater elevational change but also with greater slope change within and among pixels. The net result may be little effective signal of slope on most of the landscape occupied by forest. We explore this possibility next. While there is no generally accepted measure of spatial roughness for an areal unit, a measure of local variation proposed by Philip and Watson (1986) is widely used. This measure is computed through automatic triangulation at locations within the areal unit. Larger values of this index indicate greater roughness. In Table 5, we

Table 5
Roughness comparison between North and South regions

Quantiles					
Region	.025	.25	.5	.75	.975
North (21466)	0	.0002	.0033	.0072	.0403
South (22950)	0	0	.0018	.0061	.0369

summarize the distribution of this index over 21,466 pixels in the North region and 22,950 pixels in the South region. The table offers descriptive comparison since there is no stochastic modeling of this index. However, the quantiles for the distribution of roughness are somewhat larger in the North than the South providing quantitative support for a somewhat rougher topography.

11. Conclusions

To summarize, we have provided a new framework to explain deforestation patterns in the presence of misalignment of data layers, missing data arising from cloud cover, and incorporating spatially explicit structure through the use of a bayesian hierarchical model. Although illustrated in the context of Madagascar, the methods are more generally applicable to other ecological processes as we have noted at the end of Sections 1 and 2. Returning to an examination of population block effects, these are negatively associated spatially with forested blocks, although the effect is fairly weak. This weak signal may reflect the pervasive influence of slash-and-burn agriculture practiced throughout the landscape even areas with low population densities. This agricultural practice may be more determined by accessibility than proximity to population centers (Oxby, 1985). Elevation obviously plays into this, as do roads of any sort, or indeed even footpaths. These transportation networks are negatively associated spatially with forest blocks and positively associated with inhabited blocks. Many of these routes of transport have been used for centuries. In fact the paths followed by Ellis in the mid-1800s continue as access routes used today. In sum elevation along with road/path networks, and to a lesser extent population patterns play an important role in explaining the spatial distribution of forested blocks in the landscape. But in addition historical patterns of land use continue to play a pervasive role in the distribution of population centers and forested patches observed today in the landscape and appear to account for the differences seen between the northern and southern sections of the province.

Generalizing from our specific results, it is evident that the structural hierarchical modeling style we have employed is applicable to the modeling of land-use patterns in many other contexts. Similarity in land use between neighboring areal units would be expected, and spatial random effects could be used to capture such association. Misalignment problems among data layers, hence, between response and explanatory variables are also likely to be common themes.

Our present application has led to a model for land use, which is static, since we lack temporal information. However, were data available across time, we could extend the modeling in (1)–(4). Formally, we need only add a subscript t to those measurements which change over time. Mechanistically, we might think of the land-use process as evolving in both space and time. Spatio-temporal random effects, δ_{it} could be introduced to capture association across both space and time.

Lastly, while the present setting has an ordinal categorical response variable, in other applications the response could be binary, e.g., presence or absence of a species, or a count, e.g., abundance of a species. The first stage model for the response would change to reflect this but hierarchical modeling with spatial structure could still be employed and would provide the same benefit in terms of richer inference than is available with standard methods.

Acknowledgements

This project was supported by grants from NSF (DMS 99-71206) to A.E. Gelfand and from the John T. and Catherine D. MacArthur Foundation to R.E. Dewar and J.A. Silan-der, Jr. Roland de Gouvenain, Steven I. Citron-Pousty, Daniel L. Civco kindly provided assistance on some of the GIS component analyses of this study. Thomas H. Meyer provided feedback and discussion on spatial indices of roughness, and implemented software for estimating the index we use here. Andrew Latimer, Robin Chazdon and Richard Primack provided critical comments on the manuscript.

Detail on prior specifications

To complete the specification of the hierarchical model we require priors for α, γ, β, γ_2, γ_3 and τ^2. To obtain a well-behaved chain, we use priors for α and β that are "data-centered" with large variances. Inference is not sensitive to this centering (we could use say mean 0 centering) but our Markov chain Monte-Carlo algorithm burns in more rapidly with such centering. In particular, we assume $\beta \sim N(\hat{\beta}, k\mathbf{D}_\beta)$ where $\hat{\beta}$ is a point estimate obtained from an expectation maximization algorithm (Dempster et al., 1977), which imputes population at the pixel level and $\mathbf{D}_\beta$ is the corresponding dispersion matrix. For α, we adopt $N(\hat{\alpha}, k\mathbf{D}_\alpha)$ obtaining the estimates by fitting a multinomial ordinal response model (using standard statistical packages like SAS) that areally allocates the town population to each pixel within the town. We experimented with several values of k and found little sensitivity for $k > 10$. For the two unknown cut-points γ_2 and γ_3, we assume a uniform prior subject to the constraint $0 < \gamma_2 < \gamma_3 < \infty$. For τ^2 we adopt an inverse gamma prior. In particular $\tau^2 \sim IG(2, .23)$. This specification has infinite variance with mean roughly the sample variability in the log $\hat{\lambda}_i$ (where $\hat{\lambda}_i = P_{i.}$). As is customary, to ensure identifiability, i.e., a well-behaved posterior distribution, we impose the constraints $\sum_i \delta_i = 0$. (Besag et al., 1995).

Brief detail on model fitting

We briefly describe the conditional distributions of each variable and how they get updated (simulated) in the gibbs sampler. At time t we denote this operation by *gibbsup-date* $(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1})$ where $\boldsymbol{\theta}$ is a vector of all model unknowns that are to be updated.

- We assume the W's have a normal distribution with mean given by (5) and variance 1. Then $W_{ij} \sim N(E(W_{ij}), 1)1(\gamma_{L_{ij}-1} < W_{ij} \leq \gamma_{L_{ij}})$ and gets updated using draws from univariate truncated normals.
- $\gamma_l \sim$ Uniform(a, b); $a = \max(\gamma_{l-1}, \max(W_{ij}$ s.t. $L_{ij} = l))$; $b = \min(\gamma_{l+1}, \min(W_{ij}$ s.t. $L_{ij} = l+1))$, $l = 2, 3, \ldots$
- $\alpha \sim N((X'\mathbf{X} + D_\alpha^{-1}/k)^{-1}(X'W + D_\alpha^{-1}\hat{\alpha}/k),$ $(X'\mathbf{X} + D_\alpha^{-1}/k)^{-1})$ where $\mathbf{X}$ is the design matrix that results from the linear regression in (5). Sampling from a multivariate distribution is standard.
- $P_i \sim$ multinomial$(P_i, \lambda_{ij}/\lambda_i))\prod_j \exp(-.5(e_{ij} - \alpha_4 P_{ij})^2)$ where $e_{ij} = W_{ij} - \alpha_0 - \alpha_1 E_{ij} - \alpha_2 S_{ij} - \alpha_3 R_{ij}$, $i = 1, \ldots, T$ where T is the number of townships. We update P_i using a Metropolis step with proposal density which is multinomial $(P_i, \lambda^*_{ij}/\lambda^*_{i.})$ where $\lambda^*_{ij} = \lambda_{ij} \exp(\alpha_4 e_{ij})$. We have found this to work very well with acceptance rate of around 50%.
- $\delta_i \propto \exp(-\lambda_i)\lambda_i^{P_i} N(\sum_j \omega_{ij}\delta_j / \sum_j \omega_{ij}, \tau^2 / \sum_j \omega_{ij})$, $i = 1, \ldots, T$. This is log concave and updated using adaptive rejection sampling (Gilks and Wild, 1992). Code is available from http://www.mrc-bsu.cam.ac.uk/BSUsite/Research/ars.shtml
- $\beta_l \propto (\prod_i \prod_j \exp(-\lambda_{ij})\lambda_{ij}^{P_{ij}}) N(\hat{\beta}, k\mathbf{D}_\beta)$, $l = 0, 1, 2, 3$. This is also log concave and up dated using adaptive rejection sampling. Take care with regard to underflows when evaluating the log density.

  $\tau^2 \sim$ Inverse gamma (shape $= T/2 + 2$, scale $= .23 + \sum_{i-j}(\delta_i - \delta_j)^2/2$) where the relation i–j means i and j are neighbors. This is easily sampled using standard routines for simulating from the gamma distribution (see Devroye, 1986).

Psuedo code to fit model HI

- Initialize $\boldsymbol{\theta}$ to $\boldsymbol{\theta}_0$. We set $P_{ij0} = P_i/n_i$ and taking α_0, γ_{20}, γ_{30} estimates obtained from fitting a multinomial ordinal response model and β_0 estimates obtained by fitting Poisson regression assuming $P = P_0$. The other parameters would be automatically initialized by the operation *gibbsupdate*$(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)$.
- for i in $1{:}(B + \text{Th} \times \text{Nsamp})$ *gibbsupdate*$(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i)$
- Discard the first B iterates, store every Thth iterate thereafter and use the Nsamp iterates to make inference

  For the current application, $B = 20000$, Th $= 50$ and Nsamp $= 1000$ were appropriate.

Details on the multiple imputation

The imputation for all of the pixels with missing land-use classification is carried out in an iterative fashion using a joint spatial Potts distribution. Recall that L_{ij} denotes the classification for the jth pixel in the ith town and L denotes the set of all L_{ij}. The joint density of L under the Potts model is

$$f(L) \propto \exp(\phi > U(L)) \tag{C.1}$$

where $U(L) = \Sigma 1(L_{ij} = L_{i'j'})$ and the summation is over pairs of neighboring contiguous pixels and ϕ is an association parameter. That is, $\phi = 0$ indicates no spatial association while $\phi > 0$ encourages classification agreement between adjacent pixels. Hence the conditional distribution of L_{ij} given all of the other pixels

$$f(L_{ij}|\text{remaining } L's) \propto \exp(\phi \sum 1(L_{ij} = L_{i'j'}) \tag{C.2}$$

where the summation is over the neighbors of L_{ij}. In other words, L_{ij} takes one of the values 1, 2, 3 or 4. From (C.2), the conditional probabilities of these values reflect the relative agreement each one has with the values of the neighbors of L_{ij}.

The iterative updating proceeds as follows: Given ϕ, let $L^{(0)}$ denote all of the observed land-use classifications with a random assignment of initial values for all of the unobserved pixels. We then update in any convenient sequence, each missing L_{ij} to $L_{ij}^{(1)}$ using (C.2) where the remaining L's are fixed at current levels. An iteration is completed after each missing L_{ij} is updated. We ran 10,000 iterations keeping the final set of iterated values as one imputation for the missing L_{ij}'s. To obtain a new imputation, we can either restart the iteration at a new random $L^{(0)}$ or else continue iterating for another 10,000 iterations retaining the values at that stage.

Finally, what value of ϕ is appropriate? We can estimate ϕ using the observed L's by maximum likelihood. First we note that maximum likelihood estimation of ϕ cannot be obtained explicitly due to the awkward form of the normalizing constant involved in the likelihood. The likelihood equation obtained from (C.1) is $E(U|\phi) - U(L) = 0$. However, the left hand side of this equation is an increasing function of ϕ, so we can use a bisection algorithm to obtain the unique MLE for ϕ. For a given value of ϕ, $E(U|\phi)$ is computed using Markov Chain Monte-Carlo methods. Completely documented code written in C is available from the authors upon request. We obtained MLE's separately for the North and South by extracting a single contiguous subregion for each region with no missing land-use classification. MLE's for the northern and southern regions were 1.22 and 1.07, respectively.

References

Agarwal, D.K., Gelfand, A.E., Silander Jr., J.A., 2002. Investigating tropical deforestation using two stage spatially misaligned regression models. J. Agric. Biol. Environ. Stat. 7, 420–439.

Albert, J., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. J. Am. Stat. Assoc. 88, 669–679.

Angelsen, A., Kaimowitz, D., 1999. Rethinking causes of deforestation: lessons from economic models. World Bank Res. Observe 14, 73–98.

Apan, A.A., 1999. GIS Applications in Tropical Forestry. University of Southern Queensland Press, Toowoomba, Australia.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2003. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall.

Barbier, E.B., 2001. The economics of tropical deforestation and land use: and introduction to the special issue. Land Econ. 77, 155–171.

Benoit de Coignac, M., Bailly, M., Malvos, M., Hueber, M., Ramanahadray, M., 1973. Essai d'aménagement de terres dans la zone forestière de l'est de Madagascar. Revue Bois et Forêts des Tropiques 152, 3–18.

Besag, J., 1974. Spatial interaction and the analysis of lattice systems (with discussion). J. Royal Stat. Soc. B 36, 192–236.

Besag, J., Green, P.J., Higdon, D., Mengersen, K., 1995. Bayesian computation and stochastic systems (with discussion). Stat. Sci. 10, 3–66.

Bloom, A., 1998. A critical analysis of three approaches to tropical forest conservation based on experience in the Sangha Region. Yale Sch. Forestry Environ. Sci. Bull. 102, 208–215.

Borsuk, M.E., Higdon, D., Stow, C.A., 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecol. Model. 143 (3), 165–181.

Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecol. Model. 173 (2–3), 219–239.

Brown, K., Pearce, D.W., 1994. The Causes of Tropical Deforestation. UBC Press, Vancouver, Canada.

Campbell, G., 1991. The state and pre-colonial demographic history: the case of nineteenth-century Madagascar. J. Afircan Hist. 32, 415–445.

Carlin, B.P., Louis, T.A., 2000. Bayes and Empirical Bayes Methods for Data Analysis, second ed. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Chomitz, K.M., Gray, D.A., 1996. Roads, land use, and deforestation: a spatial model applied to Belize. World Bank Econ. Rev. 10 (3), 487–512.

Clark, J.S., 2003. Uncertainty and variability in demography and population growth: A hierarchical approach. Ecology 84 (6), 1370–1381.

Clark, J.S., LaDeau, S., Ibanez, I., 2004. Fecundity of trees and the colonization-competition hypothesis. Ecol. Monogr. 74 (3), 415–442.

Cressie, N.A.C., 1993. Statistics for Spatial Data. Wiley, New York, USA.

Dandoy, G., 1973. Terroirs et économies villageoises de la région de Vavatenina (côte orientale malagache). Atlas des structures agraires à Madagascar, 1. OSTROM, Paris, France.

de Martonne, E., 1911. La densité de la population à Madagascar. Annales de Géographic 20, 77–85.

Dempster, A., Laird, N., Rubin, D.B., 1977. Maximum likelihood prom incomplete data via the EM algorithm (with discussion). J. Royal Stat. Soc. B 39, 1–38.

Dewar, R.E., 1997. Does it matter that Madagascar is an island? Hum. Ecol. 25, 481–487.

Devroye, L., 1986. Non-Uniform Random Variate Generation. Springer-Verlag, New York.

Domenichini-Ramiaramanana, B., 1988. Madagascar. In: El Fasi, M. (Ed.), General History of Africa. III. Africa from the Seventh to the Eleventh Century. University of California Press, Berkeley, California, USA.

Duffy, S.B., Corson, M.S., Grant, W.E., 2001. Simulating land-use decisions in the La Amistad biosphere reserve buffer zone in Costa Rica and Panama. Ecol. Model. 140 (1–2), 9–29.

Du Puy, D.J., Moat, J., 1999. Vegetation mapping and biodiversity in Madagascar geographic information systems. In: Timberlake, J., Kativu, S. (Eds.), African Plants: Biodiversity Taxonomy and Uses. Royal Botanic Gardens, Kew, UK.

Ellis, W., 1858. Three Visits to Madagascar—1853–1854–1856. John Murray, London, UK.

Elster, C., 2000. Reasons for reforestation success and failure with three mangrove species in Colombia. Forest Ecol. Manage. 131, 201–214.

Esoavelomandroso, M., 1985. Madagascar 1880s–1930s: African initiatives and reaction to colonial conquest and domination. In: Boahen, A.A. (Ed.), General History of Africa. VII. Africa Under Colonial Domination, 1880–1935. University of California Press, Berkeley, California, USA.

Faramalala, M.H., 1995. Formations végétales et domaine forestier national de Madagascar. Centre National de Recherches sur L'Environnement and Conservation International, Antananarivo, Madagascar.

Flacourt, E., 1658. Histoire de la Grand Isle Madagascar avec une Relacion de ce qui s'est Passé en Années 1655, 1656 and 1657 non Encore vue par la Premierè. Chez Pierre Bien-fait, Paris, France.

Food and Agriculture Organization (FAO), 2001. The global forest resources assessment 2000 summary report. COFO-2001/INF.5.

Gelfand, A.E., Schmidt, A.M., Wu, S., Silander Jr., J.A., Latimer, A., Rebelo, A.G., 2005. Modelling species diversity through species level hierarchical modeling. J. R. Stat. Soc. Sect. C Appl. Stat. 54 (1), 1–20.

Gelfand, A.E., Silander Jr., J.A., Wu, S., Latimer, A., Lewis, P., Rebelo, A.G., Holder, M., 2003. Explaining species distribution patterns through hierarchical modeling. Bayesian Anal. 1 (1), 1–35.

Gelfand, A.E., Ghosh, S.K., 1998. Model choice: a minimum posterior predictive loss approach. Biometrika 85, 1–11.

Gelfand, A.E., Smith, A.F.M., 1990. Sampling based approaches to calculating marginal densities. J. Am. Stat. Assoc. 85, 398–409.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte-Carlo in Practice. Chapman and Hall, London, UK.

Gilks, W.R., Wild, P., 1992. Adaptive rejection sampling for Gibbs sampling. Appl. Stat. 41, 337–348.

Gourou, P., 1945. La population de Madagascar. Annales de Géographie 53–54 (296), 299–301.

Grandidier, A., Froidevaux, H., Grandidier, G. (Eds.), 1920. Mémoires sur l'ile de Madagascar par François Martin (1665–1668). Collection des ouvrages anciens concernant Madagascar, vol. 9. Union Coloniale, Paris, France, pp. 429–633.

Grandidier, G. (Ed.), 1934. Atlas des Colonies Françaises, Protectorats et Territoires Sous Mandat de la France, Société d'Éditions Geógraphiques,. Maritimes et Coloniales, Paris, France.

Granger, A., 1998. Modeling tropical land use change and deforestation. In: Godsmith, F.B. (Ed.), Tropical Rain Forest: A Wider Perspective. Chapman and Hall, London, UK.

Green, G.M., Sussman, R.W., 1990. Deforestation history of the eastern rain forests of Madagascar from satellite images. Science 248, 212–215.

Green, P.J., Richardson, S., 2002. Hidden Markov models and disease mapping. J. Am. Stat. Assoc. 97, 1055–1070.

Guariguata, M.R., Ostertag, R., 2001. Neotropical secondary forest succession: changes in structure and functional characteristics. Forest Ecol. Manage. 148, 185–206.

Hébert, J.C., 1980. Notes sur Madagascar (Côte Est et Imerina) de traitant Jacques de la Salle, recueilliés par d'Unienville en 1816. Omaly sy Anio 11, 129–175.

Hornac, J., 1943–1944. Le déboisement et la politique forestière à Madagascar. Mémoire de Stage. Mémoires de l'Ecole Coloniale ENFOM. Archives D'Outre Mer. Archives Nationales de Prance, Aix-en-Provence, France.

Humbert, H., Cours Darne, G., (Eds.), 1965. Notice de la Carte Madagascar. Extrait des Travaux de la Section Scientific et Technique de l'lnstitut Français de Pondichery, Pondichery, India.

Humbert, H., Cours Darne, G. et al. 1964–1965. Carte intemationale du tapis végétal et des conditions écologique. Trois coupures au 1 000 000^{e}. Centre National de la Recherche Scientifique and l'Office de la Recherche Scientific et Technique Outre-Mer, L'lnstitut Français de Pondichery, Pondichery, India.

Irwin, E.G., Geoghegan, J., 2001. Theory, data, methods: developing spatially explicit economic models of land use change. Agric. Ecosyst. Environ. 85, 7–23.

IUFRO, 2001. XXI IUFRO (International Union of Forestry Research Organizations) World Congress, vol. 4, Congress Report. Forestry Research Institute, Kuala Lumpur, Malaysia.

Jarosz, J., 1993. Defining and explaining tropical deforestation: shifting cultivation and population growth in colonial Madagascar (1896–1940). Econ. Geography 69, 366–379.

Kent, R.K., 1992. Madagascar and the islands of the Indian Ocean. In: Ogot, B.A. (Ed.), General History of Africa. V. Africa from the Sixteenth to the Eighteenth Century. University of California Press, Berkeley, California, USA.

Kok, K., Winograd, M., 2002. Modelling land-use change for Central America, with special reference to the impact of hurricane Mitch. Ecol. Model. 149 (l–2), 53–69.

Lamb, D., Parrotta, J., Keenan, R., Tucker, N., 1997. Rejoining habitat remnants: restoring degraded rain forest lands. In: Laurance, W.F., Bierregaard Jr., R.O. (Eds.), Tropical Forest Remnants—Ecology, Management, and Conservation of Fragmented Communities. University of Chicago Press, Chicago, Illinois, USA.

Leach, G., Mearns, R., 1988. Beyond the Fuelwood Crisis: People, Land and Trees in Africa. Earthscan, London, UK.

Le Bourdiec, F., Battistini, R., Le Bourdiec, P. (Eds.), 1969. Atlas de Madagascar. Association des Geographes de Madagascar. Bureau pour le Developpement de la Production Agricole, Tananarive, Madagascar.

Matthews, E., Payne, R., Rohweder, M., Murray, S., 2000. Pilot Analysis of Global Ecosystems: Forest Ecosystems. World Resources Institute, Washington, DC, USA.

Mercier, J.-R., 1991. La deforestation en Afrique: situation et perspectives. Edisud, Aix-en-Provence, France.

Mertens, B., Lambin, E.F., 2000. Land-cover-change trajectories in southern Cameroon. Ann. Assoc. Am. Cartographers 90 (3), 467–494.

Meyers, N., Mittermeier, R.A., Mittermeier, C.G., de Fonesca, G.A.B., Kent, J., 2000. Biodiversitiy hotspots for conservation priorities. Nature 403, 853–858.

Mugglin, A.S., Carlin, B.P., 1998. Heirarchical modeling in geographic information systems: population interpolation over incompatible zones. J. Agric. Biol. Environ. Stat. 3, 111–130.

Mugglin, A.S., Carlin, B.P., Gelfand, A.E., 2000. Fully model Based Approaches for Spatially Misaligned Data. J. Am. Stat. Assoc. 95, 877–887.

Oates, J.F., 1999. Myth and Reality in the Rain Forest: How Conservation Strategies are Failing in West Africa. University of California Press, Berkeley, California, USA.

Olson, S.H., 1984. The robes of the ancestors : forests in the history of Madagascar. J. Forest Hist. 28, 174–186.

Oxby, C., 1985. Forest farmers: the transformation of land use and society in eastern Madagascar. Unasylva 37 (2), 42–51.

Palloni, A., 1992. The relation between population and deforestation: methods for drawing causal inferences from macro and micro studies. CDE Working Paper 92-14. Center for Demographic and Ecology. University Wisconsin, Madison.

Pelet, P. (Ed.), 1902. Atlas des Colonies Françaises. Librairie Armand Colin, Paris, France.

Perrier de la Bâthie, H. 1921. La vegetation malgache. Musee Colonial, Marseille, France.

Philip, G.M., Watson, D.F., 1986. A method for assessing local variation among scattered measurements. Math. Geol. 18, 759–764.

Potts, R., 1952. Some generalized order-disorder transformation. In: Proceedings of the Cambridge Philosophical Society, vol. 48, pp. 106–109.

Prato, T., 2000. Multiple attribute Bayesian analysis of adaptive ecosystem management. Ecol. Model. 133 (3), 181–193.

Qian, S.S., King, R.S., Richardson, C.J., 2003. Two statistical methods for the detection of environmental thresholds. Ecol. Model. 166 (l–2), 87–97.

Richards, J.F., Tucker, R.P., 1988. World Deforestation in the Twentieth Century. Duke University Press, Durham, North Carolina, USA.

Schafer, J.L., 1997. Analysis of Incomplete Multivariate Data. Chapman and Hall, London, UK.

Sharma, N.P., Reitbergen, S., Heimo, C.R., Patel, J.I., 1994. A Strategy for the Forest Sector in Sub-Saharan Africa. Technical Paper No. 251. Africa Technical Department Series. World Bank, Washington, DC, USA.

Silander Jr., J.A., 2000. Temperate forests: plant species biodiversity and conservation. In: Levin, S.A. (Ed.), Encyclopedia of Biodiversity, vol. 5. Academic Press, New York, USA, pp. 607–626.

Singh, K.D., 2001. Rainforest loss and change. In: Levin, S.A. (Ed.), Encyclopedia of Biodiversity, 5. Academic Press, San Diego, California, USA, pp. 25–32.

Soares-Filho, B.S., Coutinho Cerqueira, G., Lopes Pennachin, C., 2002. Dinamica-a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecol. Model. 154 (3), 217–235.

Sponsel, L.E., Headland, T.N., Bailey, R.C., 1996. Tropical Deforestation: The Human Dimension. Columbia University Press, New York, USA.

Sussman, R.W., Green, G.M., Sussman, L.K., 1994. Satellite imagery, human ecology, anthropology and deforestation in Madagascar. Hum. Ecol. 22, 333–354.

Torsten, A., 1992. Deforestation of Tropical Rainforests: Economic Causes and Impact on Developement. J.C.B. Mohr, Tubingen, Germany.

UNEP, 1998. UN System-wide Earthwatch Web site: Madagascar. http://www.unep.ch/-islands/CMI.htm, 2/03/98.

UNEP-WCMC, 2001. Forest cover and protection for Africa, 8 January 2001 http://www.unep-wcmc.org/forest/data/cdrom2/afrl.htm.

USDA Forest Service, 2000. International Forestry Programs: Africa: Madagascar, http://www.fs.ged.us/global/globe/africa/madagascar.html.

Veldkamp, A., Lambin, E.F., 2001. Predicting land-use change. Agric. Ecosyst. Environ. 85, 1–6.

Verin, P., Griveaud, P., 1968. La protection des richesses naturelles, archeologiques, et artistiques a Madagascar. Ecole National de Promotion Sociale, Univeriste de Madagascar, Antananarivo, Madagascar.

Whitmore, T.C., 1998. An Introduction to Tropical Rainforests, second ed. Oxford University Press, New York, USA.

Wikle, C., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84, 1382–1394.

Wright, H.T., Dewar, R.E., 2000. Northeast Madagascar in the XVth to XVIIth Centuries. In: Allibert, C., Rajaonarimanana (Eds.), L'extraodinaire et le quoitidien. Karthala, Paris, pp. 287–311.

Wu, J., David, J.L., 2002. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. Ecol. Model. 153, 7–26.