

Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients

Janet Franklin

Department of Geography, San Diego State University, San Diego, CA 92182, USA

Abstract: Predictive vegetation mapping can be defined as predicting the geographic distribution of the vegetation composition across a landscape from mapped environmental variables. Computerized predictive vegetation mapping is made possible by the availability of digital maps of topography and other environmental variables such as soils, geology and climate variables, and geographic information system software for manipulating these data. Especially important to predictive vegetation mapping are interpolated climatic variables related to physiological tolerances, and topographic variables, derived from digital elevation grids, related to site energy and moisture balance. Predictive vegetation mapping is founded in ecological niche theory and gradient analysis, and driven by the need to map vegetation patterns over large areas for resource conservation planning, and to predict the effects of environmental change on vegetation distributions. Predictive vegetation mapping has advanced over the past two decades especially in conjunction with the development of remote sensing-based vegetation mapping and digital geographic information analysis. A number of statistical and, more recently, machine-learning methods have been used to develop and implement predictive vegetation models.

Key words: niche, gradient model, vegetation map, geographic information system, remote sensing, digital terrain data, habitat model.

I Introduction

Until some years ago the exclusive means of storing and presenting the geographic distribution pattern of vegetation was the classical map (Brzeziecki *et al.*, 1993: 505).

Unfortunately, many consider maps to be straightforward and think of geographic information systems as simply containers for maps (Goodchild, 1994a: 2).

Use and manipulation of spatial data have increased dramatically over the past decades with the development of geographic information-processing software and hardware.

However, it has been noted that new representations of spatial data have been embraced more slowly: '... polygons and their geometry and topology are themselves artifacts of the data modeling process... The real geographic variation is complex and continuous' (Goodchild *et al.*, 1992: 90; see also Goodchild, 1987; 1988; 1994b). It is not uncommon for land-management and planning agencies, for example, to digitize vegetation maps produced using traditional photointerpretive and cartographic methods (based on a 'communication paradigm'; see DeMers, 1991) and to incorporate them into a geographic information system (GIS) as 'data' for management decisions (e.g., as if they were based on an analytic paradigm *sensu* Tobler, 1959). In this article I review methods for predictively mapping vegetation and species distributions with inductive models to produce digital biospatial data based on an analytical paradigm.

Predictive vegetation mapping is defined here as predicting the vegetation composition across a landscape from mapped environmental variables. Predictive vegetation mapping always starts with the development of some type of model, followed by the application of that model to a geographic database to produce the predictive map, a realization of the model. Predictive vegetation mapping is founded in ecological niche theory and vegetation gradient analysis, and rests on the premises that vegetation distribution can be predicted from the spatial distribution of environmental variables that correlate with or control plant distributions. Further, maps of the environmental variables or their surrogates must be available, or easier to map than the vegetation itself, in order for predictive vegetation mapping to be a practical or informative exercise. The availability of digital maps of topography and other environmental variables, such as soils, geology and climate, and GIS software for manipulating these data, has allowed the development of predictive vegetation mapping during the past 20 years. There has been an increasing use of predictive vegetation mapping for biodiversity conservation planning (reviewed by Haines-Young, 1991), ecological restoration planning (Martinez-Taberner *et al.*, 1992) and assessing the impacts of environmental change on vegetation distributions (see, for example, Palmer and Van Staden, 1992).

In this article I will discuss the development of predictive vegetation mapping especially in the context of remote sensing-based vegetation mapping. I will illustrate the dependence of predictive vegetation mapping on ecological niche theory, and show the relationship of regional-scale predictive mapping to the global-scale modelling of vegetation in relation to climate, and to animal habitat modelling. I will focus on the prediction of plant species distributions or vegetation patterns at the 'regional' scale, e.g., where the mapped extent of the predictions are generally at or within the biogeographic range of the dominant plant species. Modelling the global distribution of plant formations, functional types or biomes based on macroclimatic factors is a closely related topic that will be summarized for comparative purposes (section IV.2), but a complete review merits a separate treatment owing to its importance for exploring the potential effects of global climate change. Finally, I will review the variables most commonly used to predict vegetation distributions, with an emphasis on topographic, hydrologic and solar radiation variables derived from digital elevation data, and compare the statistical and other methods that have been used to develop the models.

My intent in this review is to summarize recent applications of predictive vegetation mapping with respect to their goals, modelling methods, data sources and outcome, and to place them in the context of contributory literature from remote sensing, ecological theory, vegetation science, climate modelling and hydrology. I also want to acknowledge the foresight of the earliest practitioners of predictive vegetation modelling who recognized

how instrumental the marriage of gradient or site modelling with digital geographic data and computers could be in changing our perception of the map from a tool for cartographic communication to a realization of a model.

II The development of predictive vegetation mapping

I believe that the first published example of what I refer to as predictive vegetation mapping was presented by Kessell (1976; 1978; 1979) in a series of articles and monographs describing the approach he called 'gradient modelling', developed for Glacier National Park (but see also Hoffer *et al.*, 1975). 'Gradient analysis attempts to describe and understand the distribution of vegetation in response to one or more environmental, resource, and/or temporal gradients (Whittaker 1973) . . . *Gradient modeling, introduced by the author, links [such] abstract space models with real space via a GIS* (Kessell 1979)' (Kessell, in press:, emphasis added). Making extensive use of field data, Kessell developed gradient models for plant and animal species and fuel properties of the vegetation, and incorporated them into what was more than a predicted map of vegetation – it was an early spatial decision-support system for fire management in wildlands. As Kessell discusses, the theoretical foundation and methodological basis for predictive vegetation mapping are provided by niche theory and gradient analysis (discussed in section IV).

At about the same time, Roger Hoffer at Purdue University (Hoffer *et al.*, 1975) and Alan Strahler at the University of California at Santa Barbara's Department of Geography, and their colleagues, were also predicting the composition of montane forests from digital topographic data using an approach that was founded in physiographic ecology and forest site modelling. Strahler *et al.* (1978: 929) noted that 'Many ecological and silvicultural studies have shown the importance of the topographic parameters of slope angle, aspect, and relative elevation in determining vegetation composition', citing, for example, Hartung and Lloyd (1969), Wickum and Wali (1974) and Strahler (1977). In an early summary of his work in this area, Strahler (1981: 19, emphasis added) stated that

... if species composition varies systematically with terrain, topographic variables . . . can be used to improve prediction of species composition through implicit or explicit use of an ecological model . . . *The technique is a prototype for a set of tools which will become increasingly important as geobased information systems [sic] using both [remotely sensed] image and collateral data develop in the coming years*

– an insightful comment in the light of subsequent developments in GIS. In section III, I will describe the development of predictive vegetation mapping in the context of remote sensing-based mapping.

Thirdly, in 1981, Elgene Box published his monograph on empirical modelling of the global distribution of plant forms from macroclimatic variables. While I have noted that my review concentrates on studies of regional vegetation patterns, global-scale efforts will be addressed in section IV.2 owing to their importance in studies of global climate change, in establishing empirical and physiological relationships between plant distributions and climatic variables, and their hierarchical relationship to regional predictive vegetation mapping.

Almost a decade after these pioneering works, a growing number of predictive vegetation mapping efforts started appearing in the literature (Table 1). Development of predictive vegetation mapping has required the integration of GIS with remote sensing (Davis and Simonett, 1991; Davis *et al.*, 1991; and see section III), spatial analytic and

Table 1 Examples of recent literature on predictive vegetation modelling and mapping (cited frequently in this article), describing the modelling method used and the environmental variables that were used as predictors in the model. Citations are listed according to their modelling methods (see Table 2 and section V)

Study	Modelling method	Environmental variables
Franklin <i>et al.</i> , 1986	Boolean	elevation, aspect
Cibula and Nyquist, 1987	Boolean	elevation, aspect
Walker <i>et al.</i> , 1992	Boolean	elevation, topographic moisture
Martinez-Taberner <i>et al.</i> , 1992	Boolean (limits of physiological tolerances)	water chemistry (to predict aquatic macrophytes)
Miller, 1986	regression	elevation, slope, aspect diversity
Ostendorf, 1993	regression	runoff (from terrain-based hydrological model)
Fels, 1994	multiple regression	elevation, slope, aspect, slope curvature
Frank, 1988	Maximum likelihood classification (MLC)	elevation, aspect, relief, slope-aspect index, satellite spectral data
Burke <i>et al.</i> , 1989	canonical correlation	elevation, slope, aspect, fetch
Franklin and Wilson, 1991	discriminant analysis, MLC	elevation, slope, aspect, satellite spectral data
Palmer and Van Staden, 1992	contingency table analysis	elevation, annual rainfall
Lowell, 1991	discriminant analysis	original vegetation type, soil, fire history, distance to forest
Fischer, 1990	Bayesian	elevation, slope, radiation, geology, soil type, snow cover, land use
Brzeziecki <i>et al.</i> , 1993	Bayesian	temperature, precipitation, elevation, slope, aspect, soil properties
Davis and Goetz, 1990	logit regression	elevation, slope, aspect, radiation, upslope catchment area, geology
Noest, 1994	logit regression	groundwater height, dune age, duration of inundation, antecedent climate, etc.
Nicholls, 1989	GLM (logit)	temperature, precipitation, elevation, lithology, topography, exposure
Austin <i>et al.</i> , 1994	GLM (logit)	temperature, precipitation
Brown, 1994	GAM and GLM (logit)	elevation, radiation, topographic moisture, snow accumulation
Mackey and Sims, 1993	MONOMAX (maximum likelihood monotonic functions)	average daily temperature of the warmest quarter
Mackey, 1994	MONOMAX	temperature, precipitation, radiation, nutrient index based on parent material
Mackey <i>et al.</i> , in press	MONOMAX	soil texture, slope, topographic moisture
Twery <i>et al.</i> , 1991	rule-based methods	slope position
D.M. Moore <i>et al.</i> , 1991	classification tree	slope, aspect, geology, hillslope position, upslope catchment area
Lees and Ritman, 1991	classification tree	slope, aspect, geology, hillslope position, upslope catchment area, satellite spectral data
Fitzgerald and Lees, 1992	neural networks	elevation, slope, aspect, geology, upslope catchment area, satellite spectral data
Payne <i>et al.</i> , 1994	genetic algorithms	slope, aspect, geology, flow length, flow accumulation, satellite spectral data

statistical tools (Goodchild, 1987; 1994b) and large field datasets (see Walker and Moore, 1988; Kessell, 1990; Leathwick and Mitchell 1992) – a technological and methodological challenge still being addressed (Faust *et al.*, 1991).

These more recent studies have been conducted over spatial scales ranging from a few (Burke *et al.*, 1989; Martinez-Taberner *et al.*, 1992; Ostendorf, 1993; Fels, 1994) to hundreds (Davis and Goetz, 1990; Fischer, 1990; Lowell, 1991; D.M. Moore *et al.*, 1991; Mackey and Sims, 1993; Brown, 1994; Payne *et al.*, 1994), and thousands (Palmer and Van Staden, 1992; Brzeziecki *et al.*, 1993; Lenihan, 1993; Ustin *et al.*, 1993; Lynn *et al.*, in press) of square kilometers. In addition to predictively mapping the distribution of species (Davis and Goetz, 1990; Lenihan, 1993; Mackey and Sims, 1993; Fels, 1994) and plant communities or assemblages (Fischer, 1990; D.M. Moore *et al.*, 1991; Palmer and Van Staden, 1992; Brzeziecki *et al.*, 1993; Brown, 1994; Fels, 1994; Lynn *et al.*, in press), other attributes of the vegetation or ecosystem have been predictively mapped, including vegetation structure or physiognomy (Mackey, 1993; 1994), vegetation succession (Lowell and Astroth, 1989; Lowell, 1991) and plant species diversity or richness (Richerson and Lum, 1980; Miller, 1986). The products of predictive mapping have been variously described as

- ‘a predictive vegetation map’ (Davis and Goetz, 1990: 70);
- ‘a predicted . . . cover type map’ (Twery *et al.*, 1991: 50);
- ‘images showing the predicted composition of forest communities’ (D.M. Moore *et al.*, 1991: 59);
- ‘a simulated map . . . where the occurrence of different vegetation types is modelled’ (Brzeziecki *et al.*, 1993: 502);
- ‘probability surfaces . . . of species dominance . . . mapped in geographical space’ (Lenihan, 1993: 670);
- ‘a spatial prediction of the potential distribution of the vegetation’ (Mackey *et al.*, 1989: 279); and
- ‘a numerical probability map of the potential natural vegetation’ (Brzeziecki *et al.*, 1993: 499).

The last two quotations illustrate an important issue in predictive vegetation mapping – when a model is calibrated using observations of vegetation composition taken from mature or ‘climax’ vegetation stands, then *potential* natural vegetation (*sensu* Kuchler, 1988) is portrayed in the predicted map (discussed by Brzeziecki *et al.*, 1993, and see Figure 1). While often the objective in predictive vegetation mapping is to portray potential vegetation patterns for the purposes of ecosystem management or restoration, or to explore vegetation – climate relationships, sometimes predictive vegetation mapping is used to interpolate among field data points in order to map actual vegetation patterns for resource inventory or management. One could argue that environmental gradients affect the distribution of potential vegetation, and that actual or existing vegetation cannot be mapped using an empirical model unless variables related to disturbance history (such as fire or grazing, see Davis and Goetz, 1990) or actual land cover (such as remotely sensed data, see Lees and Ritman, 1991) are included as predictors in the model (Figure 1). It was within this context that researchers attempting to map actual vegetation cover based on remotely sensed satellite imagery began to incorporate mapped environmental variables to improve their maps of actual vegetation cover, and this is discussed in the next section.

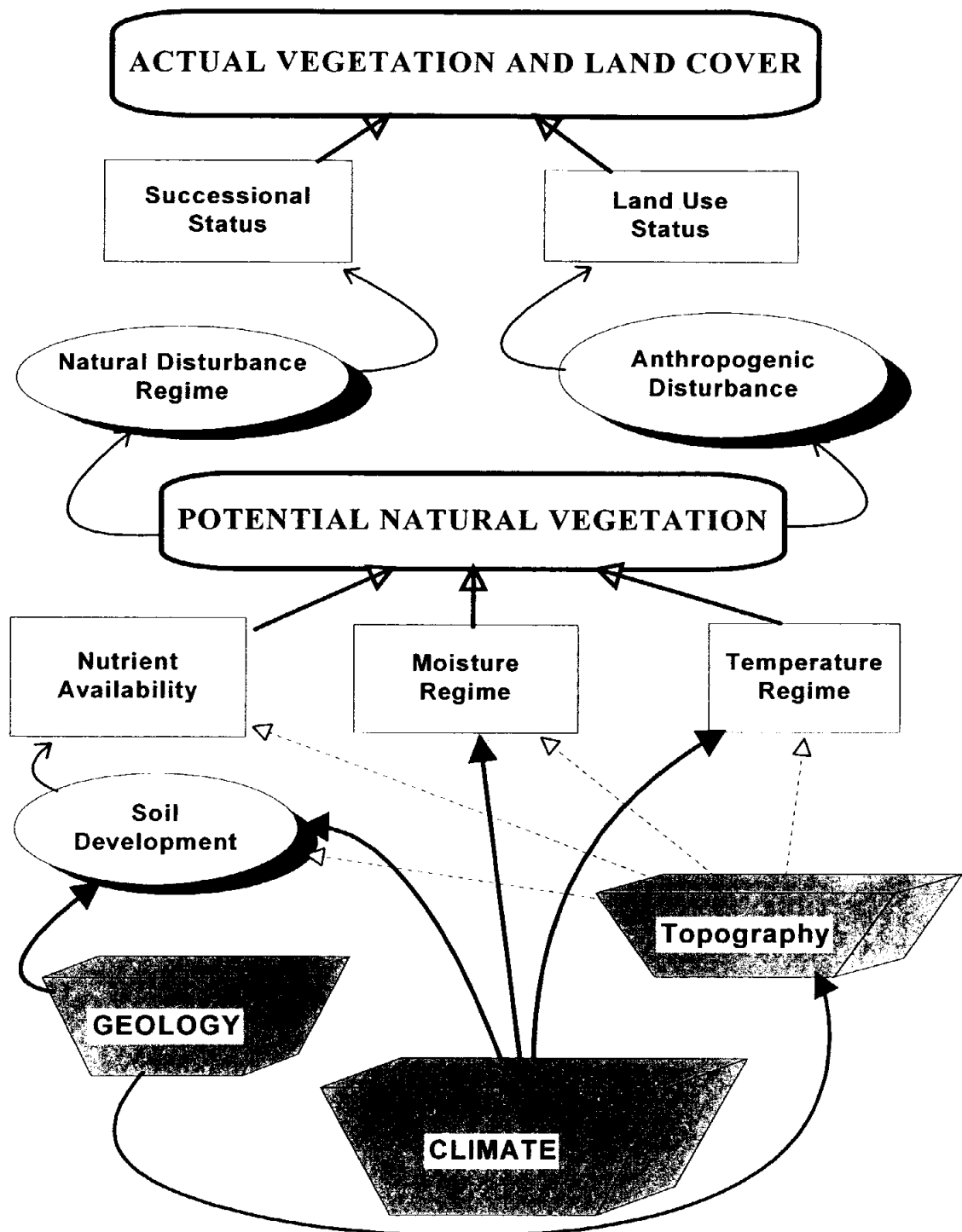


Figure 1 Conceptual model showing the relationship between direct gradients (nutrients, moisture, temperature), their environmental determinants (climate, geology, topography) and potential natural vegetation, and the processes that mediate between the potential and actual vegetation cover (the latter is sensed by a remote sensing device).

III Remote sensing and the development of digital vegetation maps

Physical geographers have been involved in analysing remotely sensed data to map earth resources since the first Landsat satellites were launched (Everett and Simonett, 1976). A number of geographers and other researchers noted early on that it was not possible to map vegetation communities in great floristic detail using the satellite multispectral imagery that was available, and they began including digital maps of environmental variables (topography, geology) in the mapping process in a number of ways (reviewed by Skidmore in his 1989a article). These environmental data were referred to as 'collateral' or 'ancillary' to the (primary) remotely sensed data (see Strahler *et al.*, 1980) and it is interesting to note in retrospect that, currently, remotely sensed data are often considered 'ancillary' to the other data in a GIS (Davis *et al.*, 1991; Ehlers *et al.*, 1991).

Charles Hutchinson (1982), a geographer, wrote an early and often-cited review of methods for combining Landsat and environmental data. At that time (and subsequently), mapping from multispectral imagery relied almost exclusively on multivariate classification techniques, identifying different land-cover types based on their spectral reflectance patterns. Hutchinson reviewed three methods for incorporating environmental data into this process:

- 'Stratification' prior to multivariate classification – dividing the study area into smaller areas (based on watershed boundaries or elevation zones, for example) to reduce the potential number of vegetation classes and separate classes that are 'spectrally similar' but geographically disjunct.
- Including ancillary data as additional variables in multivariate classification, or using them to calculate prior probabilities for vegetation classes in maximum likelihood classification (Strahler, 1980).
- After classification, dividing a 'spectral class' using some other variable (rock type, elevation) and assigning the divided class to different vegetation classes in the resulting map ('postclassification sorting').

Subsequently, terrain variables derived from digital elevation grids were included as independent variables in maximum likelihood classification or discriminant analysis to identify montane, Boreal and Arctic vegetation by Frank (1988) and S.E. Franklin and others (Franklin, 1987; Franklin *et al.*, 1989; Franklin and Moulton, 1990; Peddle and Franklin, 1991; Franklin and Wilson, 1991; 1992). S.E. Franklin (1987) refers to terrain attributes, such as slope, aspect and convexity, as 'geomorphometric variables' (see section VI) and discussed their relationship to vegetation patterns in the articles cited above.

Stratification was employed by J. Franklin *et al.* (1986) and Cibula and Nyquist (1987) in studies aimed at mapping western coniferous forest in North America. Both subdivided the study areas into regions with more or less natural bioclimatic boundaries (large watersheds or 'natural regions') and both developed simple, discrete gradient models, implemented using Boolean logic, to predict forest type from slope, aspect and elevation within each subarea (a form of 'postclassification sorting', because the broad category, conifer forest, identified using spectral classification, was subdivided into more detailed conifer associations based on the terrain variables). Satterwhite *et al.* (1984) and Shasby and Carneggie (1986) also used land-form or physiographic units to separate spectral classes 'postclassification' into vegetation classes, and in the latter case to predict vegetation type based on terrain classes. More recently, Woodcock *et al.* (1994) have implemented stratification and gradient modelling in mapping vegetation in the Sierra

Nevada, California, while developing new remote-sensing methods for delineating stands and estimating canopy structure. Other recent examples include Bolstad and Lillesand (1992), Bauer *et al.* (1994) and White *et al.* (1995). The definition of physiographic or large hydrologic units for stratification, while often implemented to assign vegetation labels to spectral classes, is also an important concept in predictive vegetation mapping when the vegetation model is based on indirect environmental gradients (see next section), and a domain for extrapolation must be defined. It may be that predictive vegetation modelling methods could be used to define 'natural regions' more objectively, based on field data.

Lees and Ritman (1991) discuss a hierarchical method of combining the complementary strengths of remotely sensed and ecological data for vegetation mapping in southeastern Australia. To paraphrase, they assert that while satellite imagery can detect land cover, and provide information about actual vegetation structure, especially when it has been affected by land-use patterns (e.g., conversion of forest to pasture; see Figure 2), in undisturbed forest, satellite imagery (broad-band multispectral observations) provides little information about gradients of species composition. However, those patterns are related to environmental gradients that are determined by topography, climate or substrate (as noted in earlier studies). They applied a tree-based rule-induction method (see section VI) calibrated with a large number of field plots and found that, as they expected, spectral variables were selected to discern anthropogenically disturbed land cover, while topographic variables were used to distinguish among undisturbed forest associations. Strahler (1981) had noted the same thing when terrain variables were combined with spectral data using maximum likelihood classification with prior probabilities.

Skidmore (1989a) applied a rule-based 'expert system' approach to mapping forest vegetation in Australia from satellite imagery and terrain data. In that study, Bayesian (probabilistic) methods were also used to adjust the prior probabilities of vegetation classes occurring in a pixel based on spectral classification, but in this case the a priori probabilities of a forest type occurring on a topographic position were elicited from expert foresters, rather than induced from a sample. These articles provide a link between the methods traditionally employed in remote-sensing image processing (parametric maximum likelihood classification) and machine-learning methods currently being explored in predictive vegetation mapping and other applications involving disparate data types and large datasets (see section VI). Ustin *et al.* (1994) developed a set of rules derived from the literature to map Boreal forest vegetation based on topographic and remotely sensed spectral variables in a deductive decision tree, analogous to many animal habitat models.

It is clear from a review of the remote-sensing literature that the use of digital environmental data to predict and study vegetation patterns has been developed in the context of remote sensing-image processing and in response to its perceived shortcomings, e.g., the inability spectrally to identify vegetation with the floristic detail desired in the resulting map. Alternatively, when maps of actual vegetation patterns could be derived from remote sensing with high accuracy, these have been compared with predictive maps generated from environmental data alone, both to validate the predictive model and to compare the distribution of potential with actual vegetation on the landscape (Davis and Goetz, 1990; Brown, 1994). Also, early on, several workers recognized the utility of digital terrain data for deriving topographic variables that have been related to vegetation patterns in many physiographic and forestry studies (section V). The relationship of species distribution patterns to environmental gradients is addressed in niche theory and discussed in the next section.

IV Gradient modelling and realized niche space

1 Plant species niches

While it is often reported (Collins *et al.*, 1993) that the continuum concept (Whittaker, 1951) underlying gradient analysis grew out of the Gleasonian (1926) individualistic hypothesis of plant species distributions, Austin (1985) maintains that the two are not synonymous. Gleason stated that no two species are alike in distribution, while the continuum concept more explicitly puts forth hypotheses about species response functions (curves) to environmental gradients, e.g., that they are Gaussian. The assumption of Gaussian response functions has not been supported by empirical evidence in Austin *et al.* (1990; 1994). Austin (1980; 1985) and Austin and Smith (1989) define three different types of environmental gradients:

- *Indirect gradients* have no direct physiological influence on plant growth; correlation with vegetation pattern is likely to be location specific: ‘. . . attempts to analyse niche relationships with . . . indirect environmental variables like slope and aspect are unlikely to be profitable, particularly when confounded with differences in location’ (Austin, 1985: 50). Indirect gradients are indicated by dashed arrows in Figure 1.
- *Direct gradients* include those having direct physiological impact but not consumed by plants (temperature, pH).
- *Resource gradients* – matter and energy used by plants for growth (light, water, nutrients, carbon dioxide, oxygen).

Austin *et al.* (1990: 161) quoted Hutchinson’s (1957) definition of species’ *fundamental niche* as ‘. . . that hypervolume defined by environmental dimensions within which that species can survive and reproduce’ or the fundamental response of plants in the absence of competition, herbivory and disease (Austin and Smith, 1989). ‘A species may be excluded from parts of its fundamental niche because of competition and other biotic interactions. The reduced hypervolume is then termed the *realized niche*’ (Austin *et al.*, 1990: 161) for which they define as a synonym ‘environmental niche’. The realized niche has also been called the species’ ecological response, realized (niche) response curve or utilization function (Austin and Smith, 1989), species response curve and realized climatic niche response (Westman, 1991), among other things, in the literature!

In gradient modelling (*sensu* Kessell, 1978) or predictive vegetation mapping, a relationship is established between environmental variables that are correlated with or surrogates for environmental or resource gradients and species distribution patterns. It follows from the above discussion that observable species distribution patterns reflect their realized rather than fundamental niches. In order to *extrapolate* over space (predictive vegetation mapping) or time (vegetation change modelling – see below), direct gradients or their surrogates must be mapped (temperature, potential solar radiation, precipitation, soil-moisture availability, geology or soil chemistry). Further, because these models are calibrated using information about the extant distribution of species (realized niche), it is problematic to predict ecosystem changes over time (e.g., in response to climate change) because biotic relationships among species are likely to change, and it would therefore be more useful to describe the fundamental (climatic) species niche (Westman, 1991; see below).

However, while Austin and Smith (1989) have noted that correlations of species distributions with indirect gradients (such as elevation) are complex and location specific,

and therefore inappropriate for modelling species' realized niches (Austin *et al.*, 1994), those relationships, if they explain sufficient variation in vegetation patterns, may be useful for *interpolating* within the environment where they were measured. D.M. Moore *et al.* (1991: 60) assert that

... while studies using direct factors have revealed much about ecological processes and relationships, it is not at all clear that such an approach is feasible for predictive modeling as it would require prior specification of the relationships between the direct gradients and the many topographic and edaphic variables that determine these.

The second approach is to model vegetation distributions directly against the topographic and edaphic variables that influence moisture and nutrient availability through such effects as drainage, exposure, and soil chemistry ... distinguishing cause from effect in such a model might be difficult ... This does not preclude the possibility of producing a predictive model for the distribution of vegetation using the many environmental variables that are correlated to it.

Other authors concur; Kessell (in press) stated that '... the abstract-space gradient models are empirical and basically predictive (rather than explanatory)', and Twery *et al.* (1991: 51) noted that '... if the rules work satisfactorily in the absence of causal understanding the goal of estimating species composition is achieved'. However, Austin *et al.* (1994) caution that the use of inappropriate statistical models (e.g., those that assume Gaussian response functions) for predictive mapping is unwise.

In plant ecology the discussion of the species niche is embedded in the community–continuum debate. Collins *et al.* (1993) suggest that viewing the vegetation as a cohesive community versus gradually changing assemblages of species distributed continuously along environmental gradients is a function of one's scale of perception. Thus the two views are complementary, rather than competitive. In an alternative interpretation, Austin and Smith (1989) maintain that the continuum is a concept referring to abstract environmental space, while the community is a spatial a concept dependent on landscape pattern: 'Co-occurring groups of species can be recognized for any particular region with a recurrent pattern of landscape' (Austin and Smith, 1989: 36). This suggests that communities (and ecotones) are geographic entities, and therefore can be predictively mapped. However, predictive mapping of species distributions presents far fewer definitional uncertainties or abstractions. Often the distribution of species assemblages (plant communities) or functional types (*sensu* Smith *et al.*, 1993) is predicted owing to methodological considerations (lack of sufficient data to model species distributions) rather than strong loyalty to the community concept.

2 Relationship of predictive vegetation mapping to models of vegetation response to climate change

There has been a resurgence of interest in climatic controls on plant distributions, motivated by a desire to predict the potential effects of rapid, anthropogenic climate change on the biosphere, and to include land–surface atmosphere feedbacks in general circulation models of the atmosphere (see Solomon, 1986; Prentice and Solomon, 1991; Henderson-Sellers, 1994). The literature on ecosystem response to global climate change links predictive vegetation mapping to other areas of ecological modelling. In a taxonomy of ecosystem models, predictive vegetation models, as they have been define here, are essentially static models (but see Lowell and Astroth, 1989; Lowell, 1991). Static models of vegetation distribution rest on the assumption that vegetation and climate are in equilibrium, or 'quasi-equilibrium' (Lenihan, 1993: 667), an assumption accepted by many for large spatial and temporal scales (Cramer and Leemans, 1993).

Traditional, empirical classification schemes of the global distribution of plant forma-

tions (biomes) as a function of climate (Köppen, 1936; Holdridge, 1947) have been refined in a number of ways (see Henderson-Sellers, 1994, for a review of the application of modified Holdridge schemes in the context of global climate model predictions). Box (1981) developed a model that predicted almost 100 plant functional types by correlation with six climatic variables. Woodward (1987) and Woodward and Williams (1987) developed a mechanistically based model, predicting the distribution of biomes from physiologically based climate variables (annual minimum temperature, for example). Prentice *et al.* (1992) predictively mapped plant functional types from a small number of physiologically based constraints related to five climate variables. Biomes were assembled from combinations of functional types. Neilson *et al.* (1992) and Lenihan and Neilson (1993) also developed rule-based models for predicting the continental-scale distribution of life forms (trees, shrubs, herbs), and biomes were defined as assemblages of different life forms. While these last three models are physiologically based, rather than correlative, they are implemented as static, rule-based, equilibrium models. The redistribution of vegetation as a function of a changed climate would be predicted solely as a function of the spatial redistribution of the driving climatic variables (although Neilson and Marks, 1994, have augmented their model with a mechanistic water-balance model that simulated plant-leaf area as a function of soil-moisture availability).

Another group of 'equilibrium models' derives a species' ecological response surface (or realized climatic niche response) from present (Westman, 1991; Lenihan, 1993; Mackey and Sims, 1993) or palaeo- (Webb, 1992) distributions on a regional to continental scale. Static or equilibrium models that address species or functional types (Smith *et al.*, 1993) are considered more powerful predictors than those predicting biome or plant community distributions owing to strong palaeoecological evidence that plant assemblages are quite transient over geologic time (Huntley and Webb, 1988). These ecological response surfaces have been derived from climate variables using species range maps (Westman, 1991; Lenihan, 1993), vegetation field-survey data (Westman, 1991; Mackey and Sims, 1993) or pollen distribution data (Webb, 1992) based on logistic regression (Lenihan, 1993) or various curve-fitting procedures (Westman, 1991; Webb, 1992; Mackey and Sims, 1993). Malanson *et al.* (1992) presented methods for estimating plant species' fundamental niche from observed distributions on environmental gradients (realized niche).

The alternative to static, equilibrium models is dynamic simulation modelling of birth, growth and death of individual plants through time '... capable of predicting the transient response of vegetation to climate change' (Lenihan, 1993: 667). The forest-stand growth and succession models (Urban *et al.*, 1991) are reviewed by Prentice *et al.* (1993). Further, a number of geographers (Hanson *et al.*, 1990; Baker *et al.*, 1991) and others (see Baker, 1989; Sklar and Costanza, 1992) are developing spatial simulation models of landscape dynamics in order to study ecosystem dynamics, and explore the effects of climate change (Baker *et al.*, 1991) and other anthropogenic modifications to natural disturbance regimes.

This brief survey of the literature on modelling changes in plant distributions in response to climate change reveals two themes. First, in a hierarchical scheme of environmental controls on plant distributions, physiologically based climatic variables that are related to direct gradients control plant distributions at the largest spatial and temporal scales (see Figure 1). Secondly, even over smaller spatial extents, predictive vegetation mapping, as I define it in this article, can serve as an equilibrium model of vegetation change when it is based on variables related to direct, rather than indirect, gradients. Neilson *et al.* (1992)

noted that developing an equilibrium model helped them to test and validate the assumptions of a dynamic model capable of simulating transient vegetation responses to environmental change.

3 Relationship to animal habitat modelling

Predictive vegetation mapping is directly and methodologically related to animal habitat modelling when the later is used to produce a predictive map of habitat suitability. Hunsaker *et al.* (1993) reviewed the role of GIS in spatial modelling of ecosystems and classified cartographic models (*sensu* Tomlin, 1990) as 'whole mosaic' models where each data layer in a GIS contains an environmental or ecological variable which is manipulated in an equation simulating an ecological response, such as animal habitat suitability (e.g., a statistical or rule-based model). Johnston (1993) also reviewed the modelling of ecological populations and communities using GIS, and noted that it is possible to model cartographically populations that are limited by environmental carrying capacity (K-strategists) based on mapped environmental variables that are related to their survival. She notes also that models for the spatial distribution of assemblages of species (communities) can be developed using GIS if statistically valid relationships exist between communities and mapped environmental variables (Johnston, 1993: 280). Because plants, not animals, are usually studied as assemblages (but see Morrison *et al.*, 1992: 246), in the later case she has essentially described predictive vegetation mapping.

Therefore, any distinction between predictive vegetation mapping and animal species habitat modelling would be somewhat artificial. A rule-based model predicting the global distribution of plant forms from physiological tolerances is, in principle, no different from a deductive (*sensu* Stoms *et al.*, 1992) habitat model for an animal species based on expert knowledge. A unified approach has been taken to the inductive modelling of plant and animal distributions or realized niches at biogeographic scales (encompassing species ranges) from climatic variables related to direct/resource gradients or physiological tolerances. This approach originated among a group of Australian researchers and has been applied to kangaroo species (Walker and Moore, 1988), introduced tree species in Africa (Booth *et al.*, 1989), two vegetation communities and two vertebrate species in Australia (Busby, 1988), Eucalyptus species in New South Wales (Austin *et al.*, 1990; 1994), rain forest in Queensland (Mackey, 1993; 1994) and Boreal tree species in Canada (Mackey and Sims, 1993).

In practice, however, the environmental factors that affects animals or micro-organisms (e.g., plant pathogens) are often different from the environmental gradients affecting plant distributions, and include vegetation composition and structure. Vegetation type is often the primary variable driving an animal habitat model both because of its direct importance for food and shelter and because it is sometimes the only available mapped habitat variable, acting as a surrogate for other habitat factors (Scott *et al.*, 1992).

V Deriving explanatory variables from digital terrain data

In a spatial hierarchy of environmental controls on vegetation distributions, climatic variables correspond to vegetation patterns at the broadest scales, followed by geology (and its effect on soil chemistry and nutrient availability) and topography, which moderates many of the macroclimatic regimes (See Figure 1). A number of studies

discussed above (sections IV.2, IV.3) address the interpolation and modelling of climatic variables related to radiation, thermal and precipitation regimes for predictive mapping (and see, for example, Hutchinson, 1987; Palmer and Van Staden, 1992; Daly *et al.*, 1994). I will limit my discussion to topographically derived variables that affect the spatial distribution of direct and resource gradients through the effect of

- elevation on temperature and precipitation;
- slope angle and aspect on radiation regime, and therefore moisture demand;
- slope angle and hillslope and drainage basin position on soil moisture, and soil development (hence, moisture-holding capacity and also nutrient availability); and
- wind exposure on temperature and moisture.

Digital topographic data can be used to model by simulation direct gradients such as potential solar radiation, to construct indices related to direct gradients (topographic moisture), and simple topographic attributes such as elevation, slope and aspect can be used as indirect gradients in model development.

In their article 'Digital terrain modeling: a review of hydrological, geomorphologic and biological applications', I.D. Moore *et al.* (1991) review the sources, availability of, quality of and data structures used for digital elevation models (DEM), and the derivation of topographic attributes and indices from them (also reviewed elsewhere: see Franklin, 1987, and Zevenbergen and Thorne, 1987, for a geomorphologic perspective; also Skidmore, 1989b; Weibel and Heller, 1991). It is important to note that DEMs frequently contain systematic and nonsystematic errors which are amplified when first and second-order differencing operations are applied to them to derive slope and aspect, for example (see Davis and Goetz, 1990). This can limit the effectiveness of predictive vegetation mapping if terrain attributes are the key predictor variables and the values of those variables derived from DEMs do not reflect field conditions.

I.D. Moore *et al.* (1991) classify topographic attributes into primary and compound. The following primary attributes were discussed as being easily derived from DEMs, and related to vegetation patterns:

- Slope or slope angle is the steepness or gradient of the slope, the rate of change of elevation, and is related to the hydrology (overland and subsurface flow velocity and runoff rate) and hence potential soil moisture and soil development (Moore *et al.*, 1993), as well as the radiation balance, of a site.
- Slope aspect or azimuth, the direction the slope faces, is related to the radiation balance, and therefore the potential evapotranspiration or 'evaporative demand' at a site.
- Specific (or upslope) catchment area (or 'drainage basin position') is related to runoff volume and variations in soil moisture (Moore *et al.*, 1988; Davis and Goetz, 1990; Quinn *et al.*, 1991; see also Marks *et al.*, 1984; Band, 1986), and soil surface properties (Moore *et al.*, 1993).
- Slope curvature, which can be decomposed into slope profile curvature, and plan or contour curvature, describes the concave- or convexness of a slope (the second derivative of elevation) and is related to subsurface water flow and hence soil-water content, and to litter accumulation and soil erosion/deposition rates which in turn are related to soil depth and texture (Fels, 1994), water-holding capacity, nutrient availability and forest-site potential (reviewed by Moran, 1982; and see McNab, 1989; 1993).

I.D. Moore *et al.* (1991: 13) describe '... analytically derived compound topographic indices' as those that can be used as surrogates for complex physical or biophysical processes, and they discuss indices of soil-water content, potential solar radiation and soil properties (as well as precipitation and erosion, not discussed here). The topographic wetness (or moisture) index of soil-water content is simply the upslope catchment area scaled by the slope and soil transmissivity; usually when calculated using a DEM the latter is unknown and set to unity (see also Jenson and Dominigue, 1988; Quinn *et al.*, 1991). Ostendorf and Reynolds (1993) calculated runoff volume by distributing runoff to adjacent grid cells based on assumptions about soil physical properties. Spatially distributed potential solar radiation has been modelled based on DEMs using energy-balance methods (in addition to Moore *et al.*'s review, see Dubayah *et al.*, 1990; Dubayah and van Katwijk, 1992; Dubayah, 1994). More complex but realistic models account for the shading from direct radiation by, as well as reflectance of radiation from, the surrounding terrain. A compound index related to soil properties is slope position or 'terrain position' (*sensu* Skidmore, 1990), defined as the relative distance between the closest stream and ridge. It is usually assigned to classes such as 'ridge, upper mid-slope, lower mid-slope, gully'. The methods for deriving slope position from a DEM are usually extensions of those methods used for automatically delineating stream networks and drainage basins (reviewed by Skidmore, 1990). Slope position is related to both the upslope catchment area and slope curvature variables, but may be better related to hillslope processes, and thus soil properties, than the former hydrological variables (Swanson *et al.*, 1988).

Almost all predictive vegetation mapping studies have used the primary terrain attributes of elevation, aspect and/or slope, derived from DEMs, as predictor variables (Table 1; Richerson and Lum, 1980; Miller, 1986; Burke *et al.*, 1989; Nicholls, 1989; Davis and Goetz, 1990; Fischer, 1990; Lees and Ritman, 1991; Lowell, 1991; D.M. Moore *et al.*, 1991; Walker *et al.*, 1992; Brzezicki *et al.*, 1993; Brown, 1994; Payne *et al.*, in press; Ustin *et al.*, 1994; Mackey *et al.*, in press), as have many wildlife habitat predictive maps (Pereira and Itami, 1991; Aspinall and Veitch, 1993). A smaller number of studies has derived other primary attributes or compound indices for prediction, including upslope catchment area (Davis and Goetz, 1990; Lees and Ritman, 1991; D.M. Moore *et al.*, 1991; Payne *et al.*, 1994), the topographic wetness index (Brown, 1994; Mackey *et al.*, in press), potential solar radiation (Davis and Goetz, 1990; Fischer, 1990; Mackey, 1993), slope curvature (Fels, 1994) and slope position (Lees and Ritman, 1991; D.M. Moore *et al.*, 1991; Twery *et al.*, 1991). Other topographically derived variables used in predictive vegetation mapping have included wind exposure or fetch (Burke *et al.*, 1989), topographic heterogeneity (Richerson and Lum, 1980) and predicted runoff from a hydrologic model (Ostendorf, 1993, Ostendorf and Reynolds, 1993). Further there is a growing literature on predictive mapping of soil properties based on their relationships with topographic variables (see Moore *et al.*, 1993).

VI Computational methods for predictive vegetation mapping

In predictive vegetation mapping, the dependent variable can be continuous (abundance or importance of a species) or categorical (presence/absence of a species, vegetation assemblage or type); and the independent variables can be continuous (rainfall, temperature, multispectral reflectance, elevation, topographic moisture, slope), grouped into ordinal classes (slope, aspect, stand age) or categorical or nominal variables (soil type, soil

properties, geologic substrate, grazing history). In this section I will summarize the computational methods used to derive inductive models and implement them for predictive mapping, although a critical review of these methods would require separate treatment.

One way of classifying the modelling methods is to divide them broadly into Boolean (discrete) methods, parametric models and machine-learning methods. Boolean methods assign a location (grid cell in a spatial database) class membership in only one class based on ranges of values of the explanatory variables (e.g., using Boolean logic or set theory). Parametric models are statistical models that predict the probability of class membership for categorical response variables, (maximum likelihood classification, general linear models) that predict the value of a continuous response variable such as species abundance, or that combine probabilities using Bayes Theorem. In this article I am defining parametric models as those that make assumptions about the underlying distribution of the data-generating process. Machine-learning methods include inductive classification (or decision) trees and genetic algorithms, as well as artificial neural networks and expert systems. These methods have been referred to as nonparametric because they make no assumptions about the underlying data distributions. They predict a probability of class membership for categorical response variables, usually based on frequency distributions in the data used to train the model.

A number of studies employed Boolean predictive mapping. Environmental measurement space was divided discretely among classes based on deductive models (discussed in section IV.2), vegetation gradient diagrams (Franklin *et al.* 1986; Walker *et al.*, 1992; Woodcock *et al.*, 1994), or ranges of observed data values (Cibula and Nyquist, 1987; Busby, 1988; Booth *et al.*, 1989).

Strahler *et al.* (1980), in their article on 'Incorporating collateral data in Landsat classification and modeling procedures', produced a table, modified from Wrigley (1979), summarizing statistical methods for combining continuous and categorical data (see Table 2). They concluded that because analysis of variance and covariance (ANOVA, MANOVA, etc.), linear regression, maximum likelihood classification (MLC) and discriminant

Table 2 Techniques for modelling with continuous and categorical data

Dependent variable	Independent Variables		
	Continuous	Mixed	Categorical
Continuous	Regression models <i>Regression tree</i> <i>GLM</i>	ANCOVA MANCOVA <i>Regression tree</i> <i>GLM</i>	ANOVA MANOVA <i>Regression tree</i> <i>GLM</i>
Categorical	MLC Logit (GLM) Discriminant analysis <i>GAM</i> <i>Classification tree</i>	MLC with priors Logit (GLM) <i>GAM</i> <i>Classification tree</i> <i>Neural networks</i> <i>Genetic algorithms</i> <i>Expert systems</i>	Contingency table Logit (GLM) <i>GAM</i> <i>Classification tree</i> <i>Neural networks</i> <i>Genetic algorithms</i> <i>Expert systems</i>

Notes: (M)AN(C)OVA: (Multivariate) Analysis of (Co-)Variance; MLC: maximum likelihood classification; GAM: general additive models; GLM: general linear models.

Source: Modified from Strahler *et al.*, 1980, who modified it from Wrigley, 1979; author's additions italicized.

analysis make assumptions about the (normal, continuous) distributions of the independent variables and models errors, equal variances and so forth, that often do not hold for the data under consideration, two methods were especially promising for modelling vegetation type (a categorical dependent variable). These were logit modelling (logistic regression) and MLC with modified prior probabilities employed in a Bayesian-type classifier (Strahler, 1980) allowing '... the mixing of parametric and nonparametric [sic] classification models for image and collateral data respectively' (Strahler *et al.*, 1980: 1020).

Once again Strahler and his colleagues foreshadowed subsequent statistical modelling efforts in this field. Several researchers have in fact predictively mapped vegetation from satellite reflectance, terrain and/or other variables using discriminant analysis when the independent variables are continuously distributed (Franklin *et al.*, 1989; Franklin and Wilson, 1991; Lowell, 1991; Peddle and Franklin, 1991). These authors noted that discriminant analysis is robust to violations of its underlying assumptions, and used the technique predictively for classifying new data, in spite of the fact that it is not usually used in this way (Strahler *et al.*, 1980). Lowell and Astroth (1989) and Palmer and Van Staden (1992) used contingency table analysis to model vegetation at widely differing scales, and Stoms *et al.* (1992) used contingency table analysis to test the sensitivity of a habitat model for the California condor. Fels (1994) examined the relationships between 27 plant species and 11 continuously distributed terrain variables using multiple regression and fit linear and quadratic models to species abundances.

A majority of probabilistic predictive modelling and mapping efforts has used logistic regression to predict species (Lenihan, 1993; Noest, 1994), vegetation assemblage (Davis and Goetz, 1990) and animal habitat (Pereira and Itami, 1991) distributions from mapped environmental variables. In a related application, Ludeke *et al.* (1990) predictively mapped the probability of deforestation for a site in Honduras from mapped environmental, spatial and socioeconomic variables using logistic regression. Although Lowell (1991) asserted that the application of logistic regression is limited to cases of two mapped classes (e.g., predicting the probability of presence/absence of a single species), in fact logit regression can be extended to the polytomous case with multiple response categories (Wrigley, 1984: 62; and see Davis and Goetz, 1990).

Austin *et al.* (1994), Austin (1985) and Nicholls (1989) introduce the statistical framework and terminology of generalized linear models (GLM) to predictive mapping. Austin (1985: 50) describes GLM as '... a less restrictive form of regression [allowing] error distributions of the dependent variable other than the normal'. Nicholls (1989: 53–54) adds that

... the constraint imposed by the assumption of a linear function between the expected data values and explanatory variables need not be accepted ... [T]he linear combination of explanatory variables ... can be set equal to what is called the linear predictor ... related to the expected data via a link function [that can include] transformations that give non-linear relationships between the linear predictor and the expectation ... where the expected data values cannot be negatively or may be logically constrained.

He gives the examples that with rank abundances the response variable could have a Poisson distribution, and with presence/absence data the binomial distribution is appropriate. For either case a log link is commonly used (e.g., a logit model). Indeed, most studies applying GLM to predictive vegetation or species mapping employ a logit model (logistic link function) as originally suggested by Strahler *et al.* (1980) and outlined by Wrigley (1984), including Austin *et al.* (1990; 1994) and Nicholls (1989) and Leathwick and Mitchell (1992) for tree species in Australia and New Zealand, and Walker and Moore (1988) for eastern gray kangaroo. Brown (1994) used a combination of general additive

and general linear models for predictively mapping Alpine vegetation in Glacier National Park in order to test hypotheses about biophysical controls on the distribution of tree-line vegetation.

Several studies have used Bayesian statistical inference to manipulate subjective or conditional probabilities, derived from remotely sensed and/or mapped environmental data, and produce probabilistic predictive maps of vegetation or animal habitat suitability. As noted above, Strahler (1980) presented a method for estimating the conditional probabilities of a vegetation type's occurrence based on topographic variables, and using those priors to modify a maximum likelihood decision rule applied to remotely sensed data, and, as a result, the estimated accuracy of his vegetation map increased from 58% to 77%. Aspinall and Veitch (1993) calculated conditional probabilities for bird species occurrence directly for spectral and topographic variables from bird-survey data and then combined these probabilities using Bayes Theorem. Pereira and Itami (1991) developed a logit model of red-squirrel habitat based on mapped environmental variables and then used Bayes Theorem to modify these probabilities with priors generated from a logistic trend surface model of habitat use. Skidmore (1989a) generated subjective a priori probabilities of forest-type occurrence on different soil types and topographic positions by interviewing expert foresters. Again, he used Bayes Theorem to integrate these prior probabilities with the probability of class membership derived from remotely sensed data using a nonparametric classifier.

When using these methods, where the probability that an observation belongs to a class is predicted based on the values of the explanatory variables, discrete class membership is usually assigned to the class with the highest probability, or exceeding some threshold. An advantage of probabilistic over Boolean methods is that the probability of class membership can be mapped, giving some quantification of the uncertainty inherent in the analytical mapping process (Goodchild *et al.*, 1992; Goodchild, 1989; 1994b). Aspinall and Veitch (1993: 537) assert that the probability values can be interpreted as an index of habitat suitability or quality in their study.

Classification and regression trees (CART – Breiman *et al.*, 1984), also referred to as decision trees (Bayes and Mackey, 1991; Lees and Ritman, 1991; D.M. Moore *et al.*, 1991; Lynn *et al.*, in press) comprise a nonparametric, probabilistic machine-learning method for inducing a set of rules to classify a categorical (classification tree) or continuous (regression tree) dependent variable based on values of the independent variables. The tree-based method described by Breiman *et al.* (1984) is binary, divisive and monothetic (Michaelsen *et al.*, 1987), and while it has been called an essentially exploratory method (Bayes and Mackey, 1991), with no formal inference procedure (Clark and Pregibon, 1992: 378), it has none the less been used predictively for vegetation mapping (Lees and Ritman, 1991; D.M. Moore *et al.*, 1991; Lynn *et al.*, in press) and ecological modelling (Michaelsen *et al.*, 1987; 1994). Michaelsen *et al.* (1987: 39) suggest that it provides a useful alternative to parametric methods when the independent variables are suspected of interacting in a nested hierarchical fashion. Tree-based methods '... can identify and express in relatively simple form non-linear and non-additive relationships ... where the relationships between the response variable and some predictor variables are conditional on the values of other predictor variables' (Michaelsen *et al.*, 1994: 675). The variables and values selected for creating the splits in the classification tree can be examined to see if they make ecological sense. (Are variables related to soil moisture and precipitation selected to separate xerophytic from mesophytic types? Do the threshold values of variables used to separate vegetation types differ on differing soil types?)

Bayes and Mackey (1991) review general linear modelling and probabilistic decision trees as methods for quantifying species – habitat response, and present an alternative model, monotonic maximum likelihood functions for relating ranked, discrete dependent variables (such as classes of rain-forest leaf size) to environmental variables. They refer to this a nonparametric model that fits a maximum likelihood function using dynamic programming and assuming only that the function is monotonic. Mackey (1993; 1994) uses this approach to produce predictive maps. This is essentially a procedure for parameterizing a non-Gaussian function related to the realized niche space of the species or functional group. An alternative approach is to generate a look up table containing probabilities for the occurrence of species or functional groups on environmental gradients (Ostendorf, 1993). This avoids any assumptions about functional relationships and allows the mapping of observations from real space into parameter space.

Other machine-learning methods include genetic algorithms and artificial neural networks. Payne *et al.* (1994) present the application of genetic algorithms, an artificial intelligence technique, used to predict both species and vegetation-type occurrence for a subset of the vegetation data discussed in Lees and Ritman (1991) and D. M. Moore *et al.* (1991). Their results were promising and they are currently testing this approach with the full dataset (K. Payne, personal communication).

Artificial neural networks are currently being applied with great enthusiasm to the classification of remotely sensed imagery, along with other spatial data, for thematic mapping (see, for example, Benediktsson *et al.*, 1993; Civco, 1993; Chen *et al.*, 1995; Foody *et al.*, 1995 and references therein), and this relates directly to my discussion of predictive vegetation mapping. Lees (1994; in press) and Fitzgerald and Lees (1992; 1994) contrast the use of decision trees versus neural nets in predictive vegetation mapping and conclude that neural nets hold great promise. They do not require as much data for model development because the vectors of variables associated with each observation used to train the network are passed through it repeatedly in the training stage (but see Benediktsson *et al.*, 1993, who assert that neural nets require a great deal of data for training). They predict the probability (or ‘possibility’) of class membership for an observation. However, a disadvantage of neural networks is that they operate as a black box, and it is not possible to examine the criteria used for classification (the ‘weights’). So spectral data, terrain variables and so forth can be put into this black box with impunity, because no assumptions are made about data distributions, and the algorithm can be used to discriminate data that are not linearly separable. However, no insights can be gained or confirmed about the functional relationship between environmental variables and vegetation distributions.

Another area of artificial intelligence that can be applied to predictive mapping is rule-based expert systems or ‘knowledge-based’ methods, where the rules or knowledge can be derived from human experts (Kenk *et al.*, 1988; Skidmore, 1989a; Twery *et al.*, 1991) or induced from empirical data (Peddle, 1995). Srinivasan and Richards (1990) discuss alternative knowledge-based techniques for combining remotely sensed and other data for land-cover mapping.

To summarize this review of modelling methods from a geographic perspective, it is expected that plant distributions on the landscape will exhibit both spatial heterogeneity and spatial dependence. However, in few cases have predictive vegetation models explicitly considered spatial autocorrelation (Goodchild, 1994b), except in trying to choose uncorrelated locations for field samples (Davis and Goetz, 1990; D.M. Moore *et al.*, 1991; Pereira and Itami, 1991). In those examples the objective was to parameterize a model

with (spatially) independent observations. But in all cases where predictive vegetation maps were produced using a model, the spatial dependence among locations where vegetation was predicted (contiguous cells in a raster grid representing the landscape) has been ignored. This spatial dependence could be exploited as a source of information, not noise, in a predictive model.

VII Summary

Predictive vegetation mapping falls within the class of static, equilibrium, whole-mosaic models of the spatial distribution of the biota. It has developed over the past two decades in concert with growing capabilities for spatial data manipulation (GIS), and advances in spatial analysis. This development has been driven by the need for large spatial vegetation datasets for environmental planning, and for basic research on the role of the biota in earth systems science.

These models, especially when based on indirect gradients, are probably most useful for interpolating vegetation – environment correlations within a region to produce spatially distributed vegetation databases. A model based on indirect gradients is limited in its ability fully to describe or test hypotheses about species niches, or simulate ecosystem dynamics; however, an equilibrium model based on direct gradients (e.g., climatically derived variables related to physiological tolerances) can be useful for understanding and predicting biogeographic patterns of species' realized niches (but cannot simulate transient ecosystem dynamics).

In either case, maps based on predictive modelling will be more powerful research, planning and pedagogic tools if they depict the uncertainty inherent in the mapping process or the probabilistic nature of many of the modelling efforts. Towards this end, exploration of the growing literature on modelling and visualizing spatial data uncertainty would be fruitful (Goodchild and Gopal, 1989; Fisher, 1991; 1994; Congalton 1994). Goodchild (1994b) asserts that any digital representation of geographic variation is necessarily an approximation of reality, and proposes a spatial statistical model of the error (unexplained variance) in digital maps, and error propagation in GIS modeling (Goodchild *et al.*, 1992). Because a number of the methods used in predictive vegetation mapping are probabilistic they could be used to parameterize an error model (Goodchild, 1994b) or to generate spatial representations of vegetation and plant distributions that incorporate their uncertainty by some other means (Woodcock and Gopal, 1995). Ostendorf (1993) noted that a map of prediction error could be used to generate hypotheses about important spatial processes not accounted for in the vegetation model.

Increasingly, machine-learning methods are used for predictive vegetation mapping, both because of the nature of data used in modelling, and the way in which environmental variables interactively affect plant distributions. However, non-Gaussian parametric models (general linear models) have also been used extensively to explore hypotheses about species niches. An important issue, especially for machine-learning methods requiring a large number of observations for model development, is how the environment is sampled to obtain the vegetation data used to calibrate the model. In many cases field data collected for other purposes are used, owing to the great expense of obtaining these data. These data may not adequately represent the distribution of vegetation along environmental gradients throughout a region. The use of digital geographic databases for efficient stratified field sampling (Michaelsen *et al.*, 1994) is a critical area of future

research in predictive vegetation mapping. Further, as noted in Section VI, few studies have employed models that explicitly account for or exploit the spatial heterogeneity and dependence inherent in biotic patterns on the physical landscape. The application of spatial statistics to predictive vegetation mapping is an area of great potential for future research.

Acknowledgements

The number of people I owe thanks for discussing these ideas, reviewing preliminary drafts of the manuscript, and providing preprints and unpublished material is so great that I will list them alphabetically: Richard Aspinall, Frank Davis, Stephen DeGloria, John Fels, Bill Fitzgerald, Arthur Getis, Michael Goodchild, Sucharita Gopal, Thomas Hack, Allen Hope, Stephen Kessell, Brian Lees, Brendan Mackey, Paul McCullough, Christine McMichael, Joel Michaelsen, Ronald Neilson, Bertram Ostendorf, Karen Payne, Serge Rey, John Stephenson, Alan Strahler, Jennifer Swenson, Mike Wells and Curtis Woodcock thank you!

References

- Aspinall, R. and Veitch, N.** 1993: Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modeling procedure in GIS. *Photogrammetric Engineering and Remote Sensing* 59, 537–43.
- Austin, M.P.** 1980: Searching for a model for vegetation analysis. *Vegetatio* 43, 11–21.
- 1985: Continuum concept, ordination methods, and niche theory. *Annual Review of Ecology and Systematics* 16, 39–61.
- Austin, M.P., Nicholls, A.O., Doherty, M.D. and Meyers, J.A.** 1994: Determining species response functions to an environmental gradient by means of a beta-function. *Journal of Vegetation Science* 5, 215–28.
- Austin, M.P., Nicholls, A.O. and Margules, C.R.** 1990: Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecological Monographs* 60, 161–77.
- Austin, M.P. and Smith, T.M.** 1989: A new model for the continuum concept. *Vegetatio* 83, 35–47.
- Baker, W.L.** 1989: A review of models of landscape change. *Landscape Ecology* 2, 111–33.
- Baker, W.L., Egbert, S.L. and Frazier, G.F.** 1991: A spatial model for studying the effects of climate change on the structure of landscapes subject to large disturbances. *Ecological Modeling* 56, 109–25.
- Band, L.E.** 1986: Topographic partitioning of watersheds with digital elevation models. *Water Resources Research* 22, 15–24.
- Bauer, M.E., Burk, T.E., Ek, A.R., Coppin, P.R., Lime, S.D., Walsh, T.A., Walters, D.K., Befort, W. and Heinzen, D.F.** 1994: Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering and Remote Sensing* 60, 287–98.
- Bayes, A.J. and Mackey, B.G.** 1991: Algorithms for monotonic functions and their application to ecological studies in vegetation science. *Ecological Modeling* 56, 135–59.
- Benediktsson, J.A., Swain, P.H. and Ersoy, O.K.** 1993: Conjugate-gradient neural networks in classification of multisource and very-high-dimensional remote sensing data. *International Journal of Remote Sensing* 14, 2883–903.
- Bolstad, P.V. and Lillesand, T.M.** 1992: Improved classification of forest vegetation in northern Wisconsin through a rule-based combination of soils, terrain, and Landsat Thematic Mapper data. *Forest Science* 38, 5–20.
- Booth, T.H., Stein, J.A., Nix, H.A. and Hutchinson, M.F.** 1989: Mapping regions climatically suitable for particular species: an example using Africa. *Forest Ecology and Management* 28, 19–31.
- Box, E.O.** 1981: *Macroclimate and plant form*. The Hague: Junk Publishers.
- Breiman, L., Friedman, J., Olshen, R. and Stone, C.** 1984: *Classification and regression trees*. Belmont, CA: Wadsworth.
- Brown, D.G.** 1994: Predicting vegetation types at treeline using topography and biophysical disturbance variables. *Journal of Vegetation Science* 5, 642–56.
- Brzeziecki, B., Kienast, F. and Wildo, O.** 1993: A simulated map of potential natural forest vegetation of Switzerland. *Journal of Vegetation Science* 4, 499–508.

- Burke, I.C., Reiners, W.A. and Olson, R.K.** 1989: Topographic control of vegetation in a mountain big sagebrush steppe. *Vegetatio* 84, 77–86.
- Busby, J.R.** 1988: Potential impacts of climate change on Australia's flora and fauna. In Pearman, G.I., editor, *Greenhouse: planning for climate change*. Melbourne: CSIRO Publications, 387–98.
- Chen, K.S., Tzeng, Y.C., Chen, C.F. and Kao, W.L.** 1995: Land-cover classification of multispectral imagery using a dynamic learning neural network. *Photogrammetric Engineering and Remote Sensing* 61, 403–408.
- Cibula, W.G. and Nyquist, M.O.** 1987: Use of topographic and climatological models in a geographic data base to improve Landsat MSS classification for Olympic National Park. *Photogrammetric Engineering and Remote Sensing* 53, 67–75.
- Civco, D.L.** 1993: Artificial neural networks for land-cover classification and mapping. *International Journal of Geographic Information Systems* 7, 173–86.
- Clark, L.A. and Pregibon, D.** 1992: Tree based models. In Chambers, J.M. and Hastie, T.J., editors, *Statistical models in S*. Pacific Grove, CA: Wadsworth, 377–419.
- Collins, S.L., Glenn, S.M. and Roberts, D.W.** 1993: The hierarchical continuum concept. *Journal of Vegetation Science* 4, 149–56.
- Congalton, R.G.**, editor, 1994: *Proceedings: International symposium on spatial accuracy of natural resource data bases: unlocking the puzzle*. Bethesda, MD: American Society for Photogrammetry and Remote Sensing.
- Cramer, W.P. and Leemans, R.** 1993: Assessing impacts of climate change on vegetation using climate classification systems. In Solomon, A.M. and Shugart, H.H., editors, *Vegetation dynamics and global change*. New York: Chapman & Hall, 190–217.
- Daly, C., Neilson, R.P. and Phillips, D.L.** 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Applied Meteorology* 33, 140–58.
- Davis, F.W. and Goetz, S.** 1990: Modeling vegetation pattern using digital terrain data. *Landscape Ecology* 4, 69–80.
- Davis, F.W., Quattrochi, D.A., Ridd, M.K., Lam, N.S.-N., Walsh, S.J., Michaelsen, J.C., Franklin, J., Stow, D.A., Johannsen, C.J. and Johnston, C.A.** 1991: Environmental analysis using integrated GIS and remotely sensed data: some research needs and priorities. *Photogrammetric Engineering and Remote Sensing* 57, 689–97.
- Davis, F.W. and Simonett, D.S.** 1991: GIS and remote sensing. In Maguire, D.J., Goodchild, M.F. and Rhind, D.W., editors, *Geographical information systems. Vol. 1. Principles*. Harlow: Longman, 191–213.
- DeMers, M.N.** 1991: Classification and purpose in automated vegetation maps. *Geographical Review* 81, 267–80.
- Dubayah, R.** 1994: Modeling a solar radiation topoclimatology for the Rio Grande river basin. *Journal of Vegetation Science* 5, 627–40.
- Dubayah, R., Dozier, J. and Davis, F.W.** 1990: Topographic distribution of clear-sky radiation over the Konza Prairie, Kansas. *Water Resources Research* 26, 679–90.
- Dubayah, R. and van Katwijk, V.** 1992: The topographic distribution of annual incoming solar radiation in the Rio Grande river basin. *Geophysical Research Letters* 19, 2231–34.
- Ehlers, M., Greenlee, D., Smith, T. and Star, J.** 1991: Integration of remote sensing and GIS: data and data access. *Photogrammetric Engineering and Remote Sensing* 57, 669–75.
- Everett, J. and Simonett, D.S.** 1976: Principles, concepts and philosophical problems. In Lintz, J.L. and Simonett, D.S., editors, *Remote sensing of environment*. Reading, MA: Addison-Wesley, 85–127.
- Faust, N.L., Anderson, W.H. and Star, J.L.** 1991: Geographic information systems and remote sensing future computing environment. *Photogrammetric Engineering and Remote Sensing* 57, 655–68.
- Fels, J.E.** 1994: Modeling and mapping potential vegetation using digital terrain data. PhD dissertation, College of Forest Resources, North Carolina State University.
- Fischer, H.S.** 1990: Simulating the distribution of plant communities in an Alpine landscape. *Coenoses* 5, 37–43.
- Fischer, P.F.** 1994: Visualization of the reliability in classified remotely sensed images. *Photogrammetric Engineering and Remote Sensing* 60, 905–10.
- Fitzgerald, R.W. and Lees, B.G.** 1992: The application of neural networks to the floristic classification of remote sensing and GIS data in complex terrain. *Proceedings of the XVII Congress ISPRS*. Bethesda, MD: American Society of Photogrammetry and Remote Sensing, 570–73.
- 1994: Spatial context and scale relationships in raster data for thematic mapping in natural systems. In Waugh, T. and Healey, R., editors, *Advances in GIS research*. London: Taylor & Francis, 462–75.
- Foody, G.M., McCulloch, M.B. and Yates, W.B.** 1995: Classification of remotely sensed data by an artificial neural network: issues related to training data characteristics. *Photogrammetric Engineering and Remote Sensing* 61, 391–401.
- Frank, T.D.** 1988: Mapping dominant vegetation communities in the Colorado Rocky Mountain front

- range with Landsat thematic mapper and digital terrain data. *Photogrammetric Engineering and Remote Sensing* 54, 1727-34.
- Franklin, J., Logan, T.L., Woodcock, C.E. and Strahler, A.H.** 1986: Coniferous forest classification and inventory using Landsat and digital terrain data. *IEEE Transactions on Geoscience and Remote Sensing* GE-24, 139-49.
- Franklin, S.E.** 1987: Terrain analysis from digital patterns in geomorphometry and Landsat MSS spectral response. *Photogrammetric Engineering and Remote Sensing* 53, 59-65.
- Franklin, S.E. and Moulton, J.E.** 1990: Variability and classification of Landsat Thematic Mapper spectral response in southwest Yukon. *Canadian Journal of Remote Sensing* 16, 2-13.
- Franklin, S.E., Peddle, D.R. and Moulton, J.E.** 1989: Spectral/geomorphometric discrimination and mapping of terrain: a study in Gros Morne National Park. *Canadian Journal of Remote Sensing* 15, 28-42.
- Franklin, S.E. and Wilson, B.A.** 1991: Vegetation mapping and change detection using SPOT MLA and Landsat imagery in Kluane National Park. *Canadian Journal of Remote Sensing* 17, 2-17.
- 1992: A three-stage classifier for remote sensing of mountain environments. *Photogrammetric Engineering and Remote Sensing* 58, 449-54.
- Gleason, H.A.** 1926: The individualistic concept of the plant association. *Bulletin of the Torrey Botanical Club* 53, 7-26.
- Goodchild, M.F.** 1987: A spatial analytical perspective on geographic information systems. *International Journal of Geographic Information Systems* 1, 327-34.
- 1988: Stepping over the line: technological constraints and the new cartography. *American Cartographer* 15, 311-19.
- 1989: Errors in objects and fields. In Goodchild, M.F. and Gopal, S., editors, *The accuracy of spatial databases*. London: Taylor & Francis, 107-13.
- 1994a: From the director: renewal. *NCGLA Update* April, 2.
- 1994b: Integrating GIS and remote sensing for vegetation analysis and modeling: methodological issues. *Journal of Vegetation Science* 5, 615-26.
- Goodchild, M.F. and Gopal, S.** 1989: *The accuracy of spatial databases*. London: Taylor & Francis.
- Goodchild, M.F., Guoqing, S. and Shiren, Y.** 1992: Development and test of an error model for categorical data. *International Journal of Geographic Information Systems* 6, 87-104.
- Haines-Young, R.** 1991: Biogeography. *Progress in Physical Geography* 15, 101-13.
- Hanson, J.S., Malanson, G.P. and Armstrong, M.P.** 1990: Landscape fragmentation and dispersal in a model of riparian forest dynamics. *Ecological Modeling* 49, 227-96.
- Hartung, R.E. and Lloyd, W.J.** 1969: Influence of aspect on forests of the Clarksville soils in Dent County, Missouri. *Journal of Forestry* 67, 178-82.
- Henderson-Sellers, A.** 1994: Global terrestrial vegetation 'prediction': the use and abuse of climate and application models. *Progress in Physical Geography* 18, 209-46.
- Hoffer, R.M. and staff.** 1975: Natural resources mapping in mountainous terrain by computer analysis of ERTS-1 satellite data. Agricultural Experiment Station Research Bulletin 919, and LARS Contract Report 061575. W. Lafayette, IN: Purdue University, 124 pp.
- Holdridge, L.R.** 1947: Determination of world formations from simple climatic data. *Science* 105, 367-68.
- Hunsaker, C.T., Nisbet, R.A., Lam, D.C.L., Browder, J.A., Baker, W.L., Turner, M.G. and Botkin, D.B.** 1993: Spatial models of ecological systems and processes: the role of GIS. In Goodchild, M.F., Parks, B.O. and Steyaert, L.T., editors, *Environmental modeling with GIS*. New York: Oxford University Press, 248-64.
- Huntley, B. and Webb, T., III** 1988: *Vegetation history*. Dordrecht: Kluwer.
- Hutchinson, C.F.** 1982: Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing* 48, 123-30.
- Hutchinson, G.E.** 1957: Concluding remarks. *Cold Springs Harbour Symposia on Quantitative Biology* 22, 415-27.
- Hutchinson, M.F.** 1987: Methods for generation of weather sequences. In Bunting, A.H., editor, *Agricultural environments: characterisation, classification and mapping*. Walingford: CAB International, 149-57.
- Jenson, S.K. and Dominigue, J.O.** 1988: Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing* 54, 1593-1600.
- Johnston, C.A.** 1993: Introduction to quantitative methods and modeling in community, population and landscape ecology. In Goodchild, M.F., Parks, B.O. and Steyaert, L.T., editors, *Environmental modeling with GIS*. New York: Oxford University Press, 276-83.
- Kenk, E., Sondheim, M. and Yee, B.** 1988: Methods for improving accuracy of Thematic Mapper ground cover classification. *Canadian Journal of Remote Sensing* 14, 17-31.
- Kessell, S.R.** 1976: Gradient modeling: a new approach to fire modeling and wilderness resource

- management. *Environmental Management* 1, 39–48.
- 1978: Forum: perspectives in fire research. *Environmental Management* 2, 291–312.
- 1979: *Gradient modeling: resource and fire management*. New York: Springer-Verlag.
- 1990: An Australian geographic information and modeling system for natural areas management. *International Journal of Geographic Information Systems* 4, 333–62.
- 1995: The integration of empirical modeling, dynamics process modeling, visualization and GIS for bushfire decision support in Australia. In Goodchild, M., Steyaert, L.L., Parks, B.O., Crane, M.P., Johnston, C.A., Maidment, D.R. and Glendenning, S., editors, *GIS and environmental modelling: progress and research issues*. Ft. Collins, CO: GIS World Inc (in press).
- Köppen, W.** 1936: Das Geographische System der Klimate. In Köppen, W. and Geiger, R., editors, *Handbuch der Klimatologie I(C)*. Berlin: Gerbruder Borntraeger.
- Kuchler, A.W.** 1988: Other technicalities. In Kuchler, A.W. and Zonneveld, I.S., editors, *Vegetation mapping*. Dordrecht: Kluwer, 157–63.
- Leathwick, J.R. and Mitchell, N.D.** 1992: Forest pattern, climate and vulcanism in central North Island, New Zealand. *Journal of Vegetation Science* 3, 603–16.
- Lees, B.G.** 1994: Decision trees, artificial neural networks, and genetic algorithms for classification of remotely sensed and ancillary data. *Seventh Australasian remote sensing conference proceedings, Melbourne, March*, 51–60.
- in press a: Sampling strategies for machine learning using GIS. In Goodchild, M., Steyaert, L.L., Parks, B.O., Crane, M.P., Johnston, C.A., Maidment, D.R. and Glendenning, S., editors, *GIS and environmental modelling: progress and research issues*. Ft. Collins, CO: GIS World Inc (in press).
- Lees, B.G. and Ritman, K.** 1991: Decision-tree and rule-induction approach to integration of remotely sensed and GIS data in mapping vegetation in disturbed or hilly environments. *Environmental Management* 15, 823–31.
- Lenihan, J.M.** 1993: Ecological response surfaces for North American Boreal tree species and their use in forest classification. *Journal of Vegetation Science* 4, 667–80.
- Lenihan, J.M. and Neilson, R.P.** 1993: A rule-based vegetation formation model for Canada. *Journal of Biogeography* 20, 615–28.
- Lowell, K.E.** 1991: Utilizing discriminant function analysis with a geographical information system to model ecological succession spatially. *International Journal of Geographic Information Systems* 5, 175–91.
- Lowell, K.E. and Astroth, J.H. jr** 1989: Vegetation succession and controlled fire in a glades ecosystem, a geographical information system approach. *International Journal of Geographic Information Systems* 3, 69–81.
- Ludeke, A.K., Maggio, R.C. and Reid, L.M.** 1990: An analysis of anthropogenic deforestation using logistic regression and GIS. *Journal of Environmental Management* 31, 247–59.
- Lynn, H., Mohler, C.L., DeGloria, S.D. and McCulloch, C.E.** In press: Error assessment in decision-tree models applied to vegetation analysis. *Landscape Ecology*.
- Mackey, B.G.** 1993: Predicting the potential distribution of rain-forest structural characteristics. *Journal of Vegetation Science* 4, 43–54.
- 1994: A spatial analysis of the environmental relations of rainforest structural types. *Journal of Biogeography* 20, 303–36.
- Mackey, B.G., Nix, H.A., Stein, J.A., Cork, S.E. and Bullen, F.T.** 1989: Assessing the representativeness of the wet tropics of Queensland world heritage property. *Biological Conservation* 50, 279–303.
- Mackey, B.G. and Sims, R.A.** 1993: A climatic analysis of selected Boreal tree species, and potential responses to global climate change. *World Resources Review* 5, 469–87.
- Mackey, B.G., Sims, R.A., Baldwin, K.A. and Moore, I.D.** in press: Spatial analysis of Boreal forest ecosystems: results from the Rinker Lake case study. In Goodchild, M., Steyaert, L.L., Parks, B.O., Crane, M.P., Johnston, C.A., Maidment, D.R. and Glendenning, S., editors, *GIS and environmental modelling: progress and research issues*. Ft. Collins, CO: GIS World Inc (in press).
- Malanson, G.P., Westman, W.E. and Yan, Y.-L.** 1992: Realized versus fundamental niche functions in a model of chaparral response to climate change. *Ecological Modeling* 64, 261–77.
- Marks, D., Dozier, J. and Frew, J.** 1984: Automated basin delineation from digital terrain data. *Geo-Processing* 2, 299–311.
- Martinez-Taberner, A., Ruiz-Perez, M., Mestre, I. and Fortenza, V.** 1992: Prediction of potential submerged vegetation in a silted coastal marsh, Albufera of Majorca, Balearic Islands. *Journal of Environmental Management* 35, 1–12.
- McNab, W.H.** 1989: Terrain shape index: quantifying effect of minor landforms on tree height. *Forest Science* 35, 91–104.
- 1993: A topographic index to quantify the effect of mesoscale landform on site productivity. *Canadian Journal of Forest Research* 23, 1100–107.
- Michaelsen, J., Davis, F.W. and Borchert, M.** 1987: A non-parametric method for analyzing hierar-

- chical relationships in ecological data. *Coenoses* 2, 39–48.
- Michaelsen, J., Schimel, D.S., Friedl, M.A., Davis, F.W. and Dubayah, R.C. 1994: Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *Journal of Vegetation Science* 5, 673–86.
- Miller, R.I. 1986: Predicting rare plant distribution patterns in the southern Appalachians of the southeastern USA. *Journal of Biogeography* 13, 293–311.
- Moore, D.M., Lees, B.G. and Davey, S.M. 1991: A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Environmental Management* 15, 59–71.
- Moore, I.D., Burch, G.J. and Mackenzie, D.H. 1988: Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. *Transactions of the American Society of Agricultural Engineers* 31, 1098–107.
- Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A. 1993: Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal* 57, 443–52.
- Moore, I.D., Grayson, R.B. and Ladson, A.R. 1991: Digital terrain modeling: a review of hydrological, geomorphologic and biological applications. *Hydrological Processes* 5, 3–30.
- Moran, M.S. 1982: Modeling timber suitability class and site index through the use of digital terrain tapes and soil maps. Masters thesis, Department of Geography, University of California at Santa Barbara.
- Morrison, M.L., Marcot, B.G. and Mannan, R.W. 1992: *Wildlife-habitat relationships*. Madison, WI: University of Wisconsin Press.
- Neilson, R.P., King, G.A. and Koerper, G. 1992: Toward a rule-based biome model. *Landscape Ecology* 7, 27–43.
- Neilson, R.P. and Marks, D. 1994: A global perspective of regional vegetation and hydrologic sensitivities from climatic change. *Journal of Vegetation Science* 5, 715–30.
- Nicholls, A.O. 1989: How to make biological surveys go further with generalized linear models. *Biological Conservation* 50, 51–75.
- Noest, V. 1994: A hydrology – vegetation interaction model for predicting the occurrence of plant species in dune slacks. *Journal of Environmental Management* 40, 119–28.
- Ostendorf, B. 1993: Spatial explicit modeling of Arctic tundra landscapes. Doctoral dissertation, Department of Biology, San Diego State University.
- Ostendorf, B. and Reynolds, J.F. 1993: Relationships between a terrain-based hydrologic model and patch-scale vegetation patterns in an arctic tundra landscape. *Landscape Ecology* 8, 229–37.
- Palmer, A.R. and Van Staden, J.M. 1992: Predicting the distribution of plant communities using annual rainfall and elevation: an example from southern Africa. *Journal of Vegetation Science* 3, 261–66.
- Payne, K., Stockwell, D. and Davey, S. A methodology for improving the accuracy of vegetation mapping using GIS, remote sensing and genetic algorithms. Paper given at the regional conference of the International Union of Geographers: 'Environment and the quality of life in central Europe: problems of transition'. 22–26 August, Prague, 1994.
- Peddle, D. 1995: Knowledge formulation for supervised evidential classification. *Photogrammetric Engineering and Remote Sensing*, 61, 409–17.
- Peddle, D.R. and Franklin, S.E. 1991: Image texture processing and data integration for surface pattern discrimination. *Photogrammetric Engineering and Remote Sensing* 57, 413–20.
- Pereira, J.M. and Itami, R.M. 1991: GIS-based habitat modeling using logistic multiple regression: a study of the Mt Graham red squirrel. *Photogrammetric Engineering and Remote Sensing* 57, 1475–86.
- Prentice, I.C., Cramer, W., Harrison, S.P., Lee-mans, R., Monserud, R.A. and Solomon, A.M. 1992: A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19, 117–34.
- Prentice, I.C., Monserud, R.A., Smith, T.M. and Emanuel, W.R. 1993: Modeling large-scale vegetation dynamics. In Solomon, A.M. and Shugart, H.H., editors, *Vegetation dynamics and global change*. New York: Chapman & Hall, 235–50.
- Prentice, I.C. and Solomon, A.M. 1991: Vegetation models and global change. In Bradley, R.S., editor, *Global changes of the past*. Boulder, CO: UCAR/Office for Interdisciplinary Earth Studies, 365–84.
- Quinn, P., Beven, K., Chevallier, P. and Planchon, O. 1991: The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. *Hydrological Processes* 5, 59–79.
- Richerson, P.J. and Lum, K.-L. 1980: Patterns of plant species diversity in California: relation to weather and topography. *The American Naturalist* 116, 504–36.
- Satterwhite, M., Rice, W. and Shipman, J. 1984: Using landform and vegetative factors to improve interpretation of Landsat imagery. *Photogrammetric Engineering and Remote Sensing* 50, 83–91.
- Scott, J.M., Davis, F., Csuti, B., Noss, R. et al. 1992: Gap analysis – a geographical approach to

- protection of biological diversity. *Wildlife Monographs* 123, 1–41.
- Shasby, M. and Carneggie, D.** 1986: Vegetation and terrain mapping in Alaska using Landsat MSS. and digital terrain data. *Photogrammetric Engineering and Remote Sensing* 52, 779–86.
- Skidmore, A.** 1989a: An expert system classifies eucalypt forest types using Thematic Mapper data and a digital terrain model. *Photogrammetric Engineering and Remote Sensing* 55, 1449–64.
- 1989b: A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *International Journal of Geographic Information Systems* 3, 323–34.
- 1990: Terrain position as mapped from a gridded digital elevation model. *International Journal of Geographic Information Systems* 4, 33–49.
- Sklar, F.H. and Costanza, R.** 1992: The development of dynamic spatial models for landscape ecology. In Turner, M.G. and Gardner, R., editors, *Quantitative methods for landscape ecology*. New York: Springer Verlag, 239–88.
- Smith, T.M., Shugart, H.H., Woodward, F.I. and Burtin, P.J.** 1993: Plant functional types. In Solomon, A.M. and Shugart, H.H., editors, *Vegetation dynamics and global change*. New York: Chapman & Hall, 272–92.
- Solomon, A.M.** 1986: Transient response of forests of CO₂-induced climate change: simulation modeling experiments in eastern North America. *Oecologia* 68, 567–79.
- Srinivasan, A. and Richards, J.A.** 1990: Knowledge-based techniques for multi-source classification: *International Journal of Remote Sensing* 11, 505–25.
- Stoms, D.M., Davis, F.W. and Cogan, C.B.** 1992: Sensitivity of wildlife habitat models to uncertainties in GIS data. *Photogrammetric Engineering and Remote Sensing* 58, 843–50.
- Strahler, A.H.** 1977: Response of woody species to site factors in Maryland, USA: evaluation of sampling plans of continuous and binary measurement techniques. *Vegetatio* 35, 1–19.
- 1980: The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment* 10, 135–63.
- 1981: Stratification of natural vegetation for forest and rangeland inventory using Landsat digital imagery and collateral data. *International Journal of Remote Sensing* 2, 15–41.
- Strahler, A.H., Logan, T.L. and Bryant, N.A.** 1978: Improving forest cover classification accuracy from Landsat by incorporating topographic information. In *Proceedings of the 12th international symposium on remote sensing of environment*. Ann Arbor, MI: Environmental Research Institute of Michigan, 927–42.
- Strahler, A.H., Estes, J.E., Maynard, P.F., Mertz, F.C. and Stow, D.A.** 1980: Incorporation of collateral data in Landsat classification and modeling procedures. In *Proceedings of the fourteenth international symposium on remote sensing of environment*. Ann Arbor, MI: Environmental Research Institute of Michigan, 1009–26.
- Swanson, F.J., Kratz, T.K., Caine, N. and Woodmansee, R.G.** 1988: Landform effects on ecosystem patterns and processes. *BioScience* 38, 92–98.
- Tobler, W.R.** 1959: Automation and cartography. *Geographical Review* 49, 526–34.
- Tomlin, C.D.** 1990: *Geographic information systems and cartographic modeling*. Englewood Cliffs, NJ: Prentice-Hall.
- Twery, M.J., Elmes, G.A. and Yuill, C.B.** 1991: Scientific exploration with an intelligent GIS: predicting species composition from topography. *AI Applications* 5, 45–53.
- Urban, D.L., Bonan, G.B., Smith, T.M. and Shugart, H.H.** 1991: Spatial applications of gap models. *Forest Ecology and Management* 42, 95–110.
- Ustin, S.L., Szeto, L.-H., Xiao, Q.-F., Hart, Q.J. and Kasischke, E.S.** 1994: Vegetation mapping of forested ecosystems in interior central Alaska. *Proceedings IGARSS '94*. Piscataway, NJ: IEEE Inc, 229–31.
- Walker, P.A. and Moore, D.M.** 1988: SIMPLE: an inductive modeling and mapping tool for spatially-oriented data. *International Journal of Geographic Information Systems* 2, 347–63.
- Walker, R.E., Stoms, D.M., Davis, F.W. and van Wagtenonk, J.** 1992: Modeling potential natural vegetation from a topographic gradient in the southern Sierra Nevada, California. *Proceedings GIS/LIS-92*. Bethesda, MD: ASPRS, 794–803.
- Webb, T.III** 1992: Past changes in vegetation and climate: lessons for the future. In Peters, R.L. and Lovejoy, T.S., editors, *Global warming and biological diversity*. New Haven, CT: Yale University Press, 59–75.
- Weibel, R. and Heller, M.** 1991: Digital terrain modelling. In Maguire, D.J., Goodchild, M.F. and Rhind, D.W., editors, *Geographical information systems. Vol. 1. Principles*. Harlow: Longman, 269–97.
- Westman, W.E.** 1991: Measuring realized niche spaces: climatic response of chaparral and coastal sage scrub. *Ecology* 72, 1678–84.
- White, J.D., Kroh, G.C. and Pinder, J.E., III** 1995: Forest mapping at Lassen Volcanic National Park, California, using Landsat TM data and a geographic information system. *Photogrammetric Engineering and Remote Sensing* 61, 299–305.
- Whittaker, R.H.** 1991: A criticism of the plant asso-

- ciation and climatic climax concepts. *Northwest Scientist* 25, 17–31.
- 1973: Direct gradient analysis: In Whittaker, R.H., editor, *Handbook of vegetation science 5: ordination and classification of communities*. The Hague: Junk Publishers, 9–50.
- Wickum, D.A. and Wali, M.K.** 1974: Analysis of a North Dakota gallery forest: vegetation in relation to topographic and soil gradients. *Ecological Monographs* 44, 441–64.
- Woodcock, C.E. and Gopal, S.** 1995: Remote sensing of forests: new data layers for GIS. In 1995 *ACSM/ASPRS annual convention and exposition technical papers*. Bethesda, MD: American Society of Photogrammetry and Remote Sensing, 420–28.
- Woodcock, C.E., Collins, J., Gopal, S., Jakabhazy, V.D., Macomber, S., Ryherd, S., Harward, V.J., Levitan, J., Wu, Y. and Warbington, R.** 1994: Mapping forest vegetation using Landsat TM imagery and a canopy reflectance model. *Remote Sensing of Environment* 50, 240–54.
- Woodward, F.I.** 1987: *Climate and plant distribution*. Cambridge: Cambridge University Press.
- Woodward, F.I. and Williams, B.G.** 1987: Climate and plant distribution at global and local scales. *Vegetatio* 69, 189–97.
- Wrigley, N.** 1979: Development of the statistical analysis of categorical data – a review. *Progress in Human Geography* 3, 315–55.
- 1984: *Categorical data analysis for geographers and environmental scientists*. London: Longman.
- Zevenbergen, L.W. and Thorne, C.R.** 1987: Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* 12, 47–56.