Current approaches to modelling the environmental niche of eucalypts: implication for management of forest biodiversity

M.P. Austin *, J.A. Meyers

C.S.I.R.O. Division of Wildlife and Ecology, P.O. Box 84, Lyneham, A.C.T. 2602, Australia

Abstract

Robust predictive models of the distribution of forest biota are important tools for the management of forest biodiversity. To build robust models, it is essential to understand the environmental processes which control species distribution and hence choose appropriate predictor variables. Requirements for modelling the environmental niche of plant species include: environmentally stratified survey data of the vegetation and associated environmental measurements, an understanding of ecological theory, robust statistical models and geographical representation of the models. These requirements can be satisfied in different ways, many of which are discussed. The choice of modelling technique and curve fitting function should be related to ecological theory. Prediction becomes increasingly robust and less location-specific as the predictor variables become more process-oriented and relevant to biological processes. However, the need to use predictors for which estimates are available for unsampled regions may limit the choice to less direct variables.

In this context we examine the performance of two modelling techniques: Generalised Linear Modelling (GLM) and Generalised Additive Modelling (GAM). Trees are ideal to study, because their size and immobility make for ease of collecting data and they provide important habitat for fauna and understorey herbs and as such are useful for predicting the distribution of some other biota. The data set includes 8377 sites in south-eastern Australia, with presence/absence data for trees and seven environmental predictors. A detailed comparison is described for Eucalyptus cypellocarpa. The influence of 'naughty noughts', or zero values beyond the range of a species, can distort the response function, giving positive predictions where the species is known to be absent. The model is improved by restricting the data to a suitable range. GAM has advantages over GLM due to the flexible nature of the non-parametric smoothing function. Different response curves can produce divergent predictions of species occurrence, particularly at the limits of species distribution. Conservation evaluation often requires making predictions in unsampled areas, and so the assumption of particular shapes of response curve could lead to significant errors in the estimation of the conservation value of these areas.

Keywords: Survey design; Environmental niche; Generalised additive model; GAM; Generalised linear model; GLM; Eucalyptus

1. Introduction

Trees determine the microclimate, nutrient availability, shelter and food supply for many of the forest biota. Knowledge of the environmental conditions under which the different tree species occur is an essential requisite for forest management.

Information on the distribution of forest biota is often of poor quality and difficult to obtain, even for tree species in well-sampled forests such as those in the temperate regions of south-eastern Australia. For example, a description of the distribution of Euca-
**lyptus globoidea** has the species occurring from sea level to 1100 m, between latitudes 30° and 38°S, with the mean maximum temperature of the hottest month in the range 22–31°C and mean annual rainfall of 650–1400 mm (Boland et al., 1984). From this description, when *E. globoidea* occurs in areas with maximum temperatures of 22°C, it is impossible to determine whether it is only found in low rainfall areas, only in high rainfall areas, or in areas across the full range of rainfall. In such descriptions, there is no information about the combinations of conditions necessary for the occurrence of species. The trees of these eucalypt forests provide a basis to study the requirements needed to predict the environmental niche of species. Trees, because of their size and immobility, are much easier to study than other forest biota. Their importance as habitat for fauna and understorey herbs means they can be used as surrogates for predicting the distribution of other biota. In addition, their ease of study means that results can provide experience and guidelines for the more difficult study of mobile, vulnerable and infrequent species.

This paper reviews the approaches adopted in studying the distribution and ecology of eucalypts in forests in temperate eastern Australia in the context of requirements for predictive rigour. A detailed case study of selected species is provided to examine current approaches to the statistical modelling of species distribution.

### 2. Requirements for prediction

Ideally, information used for management of forest ecosystems should be the best possible and obtained in a cost-effective way. The role of the scientist should be to demonstrate what is possible, and what is lost if various compromises are made. The manager can then use this information to determine how best to allocate time, effort and resources to data collection and analysis.

Requirements necessary to model the environmental niche of trees include:

1. Geocoded plot data of vegetation from an environmentally stratified survey.
2. Associated plot data for appropriate environmental predictor variables.
3. Explicit ecological theory for relating stratification variables, environmental variables and species response functions.
4. Statistical models that give confidence limits, are compatible with ecological theory, and are based on a random sample with statistically independent observations.
5. Geographical information system with suitable data layers to display the statistical models for all locations.

Each requirement can be satisfied in different ways and many compromises are usually necessary. Each one is discussed in detail below.

#### 2.1. Survey data

No survey can provide a complete inventory of a region. Instead, only a sample of the region is measured from which predictions are made across the whole region. Margules and Austin (1994) discuss the need for explicit data (i.e. at least geocoded presence/absence data and not presence-only species lists), designed surveys and statistical analyses for studies of biodiversity, extinction and conservation. They state that few areas of the world have adequate biological field surveys or comprehensive data bases and that it is essential to have a suitable conceptual framework for regional conservation evaluation.

The ideal of an unbiased stratified survey based on an adequate sampling frame is not usually possible because of cost constraints. Austin and Heyligers (1989, 1991) argue that, for most ecological purposes including environmental niche modelling, representative sampling is more important. With this approach, the range of variation in vegetation is sampled rather than trying to obtain accurate estimates of the mean value of basal area for species in the region. In a case study from northern New South Wales, Australia, Austin and Heyligers (1989) used the simple theoretical concept that vegetation varies with temperature, rainfall and lithology, and they used the environmental space defined by these variables as the sampling frame. In this way they obtained an environmentally representative sample of the floristic variation in forest composition in the region. Austin and Heyligers (1989) used a ‘grad- sect’ (gradient) sampling design, which is cost-effective but gives a spatially biased sample due to the
use of transects orientated along gradients of maximum environmental change. Ferrier (New South Wales National Parks and Wildlife Service, personal communication, 1994) and others have used the concept of environmentally representative stratification to design a survey for the entire north coast region of New South Wales without using gradients (see also Neave et al., 1996).

Usually, the data used in forest biodiversity studies are obtained by collating existing data sets from ad hoc surveys of portions of the region under study. Consequently, the resulting data base is often spatially and environmentally biased. The extent of the environmental bias can be evaluated by using the environmental space concept. Fig. 1(a) shows the environmental space for a large region of south-eastern New South Wales together with 8377 plots obtained by collating numerous surveys (Austin et al., 1990), together with the distribution of *Eucalyptus cypellocarpa* in the space. The plots appear representative of the environmental space but by locating the plots in geographical space there is evidence (Fig. 1(b)) of spatial bias in the sampling with large gaps in some areas. Such bias does not preclude modelling but should not be ignored.

### 2.2. Environmental predictor variables

Modelling success depends on using appropriate environmental predictors. There are many possible environmental variables to choose from and selection depends on insight and awareness of relevant ecological processes. There is a continuum of environmental variables from indirect variables like altitude or latitude which have no direct physiological impact on growth or survival processes, to direct causal (resource) variables such as nitrate concentration at the root surface (Austin, 1980). Prediction becomes increasingly robust and less location-specific as the predictor variables become more process-oriented and relevant to biological processes. However, the need to use predictors for which estimates are available for unsampled regions may limit the choice to less direct variables.

Environmental process models which link indirect variables such as altitude to direct variables such as temperature are used infrequently but overcome some of the difficulties associated with the lack of ‘direct’ variable data. An example of the transformation of the usual field observations of altitude, slope and aspect into a more ecologically relevant variable is the calculation of radiation values for sites. The radiation estimate is based on complex trigonometric functions relating sun position to slope and aspect, taking into account the interception of direct radiation due to the blocking effect of topography for southern slopes (in the southern hemisphere) and the contribution of diffuse radiation from the sky. The...
relationships of radiation with the indirect variables of slope, aspect and altitude are often too chaotic to show a clear signal. Numerous algorithms exist for calculating radiation inputs and for some time these have been used directly and as input to moisture indices for predicting the distribution of species (e.g. Austin and Cunningham, 1981). Mackey et al. (1988) used an improved version incorporating regional variation in radiation and changes in the relative contribution of direct and diffuse radiation to help evaluate representativeness of areas for conservation.

Moore et al. (1993) have vastly extended these environmental process models to provide estimates of ecologically relevant predictor variables. They used process knowledge of climate and hydrology in conjunction with a digital terrain model (DTM) and geographical information system (GIS) to estimate local temperature, radiation and topographical wetness. These estimates showed correlations with forest communities for an area in the Brindabella Ranges near Canberra, Australia.

Major improvements in prediction can be obtained when these environmental process models are used to estimate a relevant predictor from simple field variables. Austin (1979) (see also Austin and Cunningham, 1981) demonstrated improvements in the $r^2$ of regression equations from 0.1 to 0.6 by substituting radiation and moisture indices for slope, aspect and soil depth.

2.3. Ecological theory

Use of microclimatological and hydrological theory to estimate environmental predictors requires complementary ecological theory to define the expected response function of organisms. Niche theory provides a number of hypotheses about the expected response of species to environmental variables. The traditional assumption is the Gaussian response curve. The fundamental niche (i.e. the species response without competition or other biotic influences) has a broad Gaussian curve. This curve is narrowed by competition to give the realised niche, also a Gaussian curve centred on the optimum position on the environmental gradient (Giller, 1984). An alternative model exists for plants (Mueller-Dombois and Ellenberg, 1974) based on the assumption that a superior competitor may be restricted by environment but can still displace another species from its physiological optimum. The realised niche and its optimum will differ from the fundamental niche in both position of the optimum and shape. In fact the realised niche may assume various skewed shapes, even bimodal.

Austin (unpublished data, 1995) reviews these alternative niche theories and their consequences for ecological research.

In practical management terms, failure to recognise the alternative shapes possible for the realised niche response to environmental gradients has led to the adoption of inefficient methodologies for predicting species distribution and performance in relation to environment. Thus linear responses to environmental predictors are often fitted, and discriminant or other multivariate methods based on the assumption of linear relationships between variables are used. Niche theory postulates that the expected response to an environmental variable will be at least curvilinear. Recent evidence suggests that responses are predominantly skewed (Austin et al., 1990, 1994a; Leathwick and Mitchell, 1992) which means that methods based on symmetric Gaussian curves are likely to give biased predictions.

Ecological theory is needed to guide both the choice of methods and predictor variables if adequate predictive models are to be attained.

2.4. Statistical models

The majority of regression models used in ecology to predict the abundance of species are still based on linear relationships, with data assumed to have normal errors. This is incompatible with ecological theory and the data rarely have normal errors. Two statistical predictive methods have recently been promoted in the ecological literature: generalised linear modelling (GLM) and generalised additive modelling (GAM). The first allows error distributions other than normal to be used. The second is an extension of the first and allows non-parametric analysis without the necessity to specify the functional form of the relationship (i.e. whether it is linear, cubic or exponential). GLM has been used extensively in Australia to predict eucalypt species distributions (e.g. Austin and Cunningham, 1981; Austin et al., 1990, 1994a), to examine patterns of species richness (Margules et al., 1987) and for
evaluating the conservation value of reserves (Margules and Stein, 1989). Nicholls (1989, 1991) provides clear examples of the use of this technique.

Generalised additive models were introduced into ecology by Yee and Mitchell (1991) and have since been used by Norton and Mitchell (1993). The ability to leave the form of the response functions undefined is a major opportunity to examine the shape of the response without making restrictive assumptions. Austin et al. (1994a) used β-functions that can assume a variety of skewed shapes, with GLM, to show that skewed responses occurred, but the nature of the skew is still restricted by the mathematical properties of the β-function. However, it is not true to say that GAM is assumption-free (Norton and Mitchell, 1993). For example, the assumptions made in developing significance tests with non-integer degrees of freedom require assessment.

Progress in prediction of the distribution of species is dependent on the evaluation of the relevant roles of GLM and GAM. The original developers of GAM, Hastie and Tibshirani (1990), recommended that GAM be used to determine the appropriate functional forms to be used in a parametric regression model. One problem in statistical modelling is the impact of ‘naughty noughts’ on regression analysis. If observations extend beyond the feasible range of a species, zero values will be recorded for the species. If large numbers of these zero values are included in regression analysis, they will reduce the explanatory power of the model and distort the shape of the response function. In effect, the models are attempting to predict the occurrence of a species beyond the domain where it can exist. Thus, separate models should be fitted to the different regions of the environmental space in order to attain adequate models (Austin and Cunningham, 1981; Austin et al., 1994a).

3. Case study of predictive modelling of eucalypt species

A review of the entire range of modelling techniques is beyond the scope of this paper. Instead, we concentrate on the recent progress in statistical modelling and niche theory of the response function, comparing the performance of GLM and GAM. The data and models have been described extensively elsewhere (Austin et al., 1990, 1994a).

3.1. Survey data

The survey data base consists of 8377 plots for which the presence/absence of 163 tree species are recorded and seven environmental predictor variables are measured or estimated. The data have been collated from numerous surveys in the southeastern corner of New South Wales, Australia. Details of the surveys and their authors can be found in Austin et al. (1990). The study area is approximately 40 000 km² and covers a wide range of environmental conditions, with altitude 0–2200 m, rainfall 500–2126 mm (Fig. 2(b)), marked variations in rainfall seasonality (Austin et al., 1990) and a variety of lithological substrates.

3.2. Predictor variables

The seven predictor variables selected for use in modelling the distribution of tree species are based on the availability of suitable information and a simple conceptual model of the expected importance of variables. Plants respond to temperature, moisture, nutrients and light. Estimates of the appropriate proximal causal variables affecting plant performance cannot be obtained and, as such, approximations are necessary. Temperature was represented as mean annual temperature and was estimated using the ES-OCLIM method based on data from 86 weather stations (Hutchinson, 1984; Adomeit et al., 1987) and uses altitude, latitude and longitude. Moisture was represented by mean annual rainfall, rainfall
seasonality and topographical position. Mean annual rainfall was estimated in a similar way to temperature, but was based on records from 635 rainfall stations. Rainfall seasonality was represented as a categorical variable or factor with three levels; winter, aseasonal, and summer rainfall derived from a study of the rainfall climate of the region by Austin and Yapp (1978). Topographical position, which is represented as a factor with six levels—ridge, slope, lower slope, gully, flat and unknown—has a marked impact on the availability of moisture, and in the absence of a suitable DTM serves as a surrogate for a wetness index.

No suitable estimates are available of soil water content or soil nutrient status. Lithology was taken from the 1:500,000 Monaro geology map and sorted into five classes as a crude estimate of soil texture and hence soil moisture status. An estimate was made of the phosphate content of the local rocks from geochemical measurements for lithologies in the region. This was expressed as a factor with five levels. Light was estimated as total radiation received at a site. The radiation was first estimated using ESOCLIM with the four predictors: altitude, latitude, longitude and estimated mean annual rainfall. These estimates were then adjusted using a modified version of CLOUDY which took aspect and slope into account (Mackey et al., 1988).

Thus for GAM and GLM modelling of tree occurrence, three continuous derived variables—mean annual temperature, mean annual rainfall and mean monthly solar radiation—were used together with four factors—rainfall seasonality, topographical position, lithology and phosphorus status—as in Austin et al. (1994a).

3.3. Statistical model

A critical issue is whether the non-parametric GAM provides evidence of different (better?) response shapes than the parametric GLM with a $\beta$-function or other curvilinear function.

Models were developed for the nine eucalypt species used by Austin et al. (1994a). As an example, we describe the process and models derived for *E. cypellocarpa*. The statistical package S-PLUS (Statistical Sciences, Inc., 1993) was used to derive two GAM models using a non-parametric smoothing spline for the continuous variables. One model was based on the full unrestricted data set and the other had the same rainfall, temperature and seasonality restrictions as for the models published by Austin et al. (1994a). Each variable was tested in turn for significance ($P < 0.001$) and a model was built in a forward stepwise procedure. Linear and non-linear (smooth) functions were tested for the continuous variables. A backwards elimination test was also made to check the significance of the variables in the final model. These models were then compared with the $\beta$-function GLM models obtained by Austin et al. (1994a) for *E. cypellocarpa*.

3.4. GIS

The lack of a suitable digital elevation model (DEM) with sufficient resolution meant that a DTM could not be used and limited the type of prediction that could be made from the statistical models. The various data layers are based on a gridcell size of 0.01 of a degree (approx. 1 km$^2$), and each cell is
assumed to be homogeneous. Predictions based on models including topographical position estimated at a scale of 0.1 ha are not possible without some simplifying assumptions. Each cell is assumed to contain all possible topographical and aspect combinations. Each combination is assigned typical slope and aspect values for which a relative radiation index value is calculated. Prediction maps for each recognised combination can then be generated. Nine combinations are used: ridge, exposed (northern), intermediate and protected (southern) slopes, exposed and protected lower slopes and gullies plus flats.

3.5. Predictive models for Eucalyptus cypellocarpa

Fig. 1 summarises the distribution of sampled plots in both environmental (Fig. 1(a)) and geographical (Fig. 1(b)) space. The distribution of *E. cypellocarpa* clearly indicates that there are limits beyond which the species does not occur. The question is how to estimate suitable limits for restricting the data analysis to avoid the 'naughty noughts' problem. This has been done arbitrarily by setting a limit to include at least 100 observations of the species absence above and below the last presence. The limits for *E. cypellocarpa* are set at 8.75°C and 16.15°C for temperature and at 681 mm minimum for rainfall (Table 2 in Austin et al., 1994a). Only one observation occurs outside the summer rainfall or coastal zone and the data set was therefore restricted to this zone.

The restricted data set consists of 5048 out of the possible 8377 plots. A GLM model was fitted to this restricted data set with a β-function for temperature within the limits: $\beta$ function for temperature = $a\log(t - 8.75) + b\log(16.15 - t)$ where $a$ and $b$ are coefficients and $t$ is mean annual temperature. The full GLM regression model for *E. cypellocarpa* is:

1. β-function for temperature in the range of 8.75°C to 16.15°C.
2. Linear function for rainfall above 681 mm.
3. Quadratic function for radiation.
4. Coefficients for each factor level for topographical position and nutrient status.

The factor lithology was non-significant.

The regression model is a mixed model with a set of subjective restrictions based on the total data set.

![Fig. 3. Predicted probability of occurrence of *E. cypellocarpa* in environmental space using a mixed model with qualitative restrictions and a GLM for gullies at medium nutrient level and high mean annual radiation.](image)
and a GLM model within the restricted environmental space. The shape of the individual response functions for the continuous variables is shown in Fig. 2. The response to temperature is a skewed function with a maximum at 11°C and a long tail towards higher temperatures (Fig. 2(a)). The β-function curve differs significantly from a symmetric curve (Austin et al., 1994a). There is a simple negative linear response to rainfall (Fig. 2(b)) and the response to radiation is quadratic with a maximum at 13.65 kJ m⁻² per day (Fig. 2(c)). The coefficients for topographical position predict maximum occurrence in gullies and minimum on ridges. The coefficients for nutrient status indicate a positive response to increasing nutrient status. Fig. 3 shows the predicted distribution of *E. cypellocarpa* in the restricted climate space of the coastal summer rainfall zone.

Two GAM models were also fitted to the *E. cypellocarpa* presence/absence data. The first based on the entire 8377 data set and the second based on the restricted data set of 5048 plots with the same temperature limits given above. The same factors were fitted, together with the smoothing functions for the three continuous variables. The regression model is a mixed model with a set of subjective restrictions based on the total data set and a GAM model within the restricted environmental space. The results for the two GAM models do not differ markedly within the range of the restricted data set for either temperature or rainfall (Fig. 4). The displacement between the two curves is due to the inclusion of the factor rainfall seasonality in the model for the unrestricted data set. The unrestricted GAM predicts a probability of occurrence of 0.2 for temperatures beyond the observed limits of the species (Fig. 4(a)). There are consistent differences in the shape of the temperature response between the GAM models and the GLM. A similar effect can be seen for rainfall (Fig. 4(b)), while the responses to radiation show a difference in shape between the GLM and GAM models at low radiation levels (Fig. 4(c)).

These predictive models can be translated into a spatial prediction using a GIS (Fig. 5). The predicted distribution of *E. cypellocarpa* on protected gullies (Fig. 5(a)) and exposed northern slopes (Fig. 5(b)), shows the major contributions that topographical position and radiation can make to the model.

4. Comparison of GAM and GLM models

GAM models have been developed for the nine eucalypt species modelled using GLM by Austin et al. (1994a). The following features are apparent from comparisons of these models (Figs. 4, 6 and 7).

1. The unrestricted GAM frequently predicts a large probability of occurrence beyond the observed limits of the species.
2. The restricted and unrestricted GAM models correspond closely within the range of restricted data.
3. Although flexible, the β-function often fails to match the shape predicted by the GAM models.
4. The GAM models for temperature are consistent with the skewed curves obtained using GLM.

The discrepancy between the β-function GLM model and the GAM models for describing the temperature response is noticeable. For *E. cypellocarpa* (Fig. 4(a)), GAM models suggest there may be an additional point of inflexion in the skewed response. For *Eucalyptus rossii* (Fig. 6(a)) and *Eucalyptus fastigata* (Fig. 6(b)), the results indicate that the
Fig. 5. Geographical distribution of *E. cypellocarpa* predicted for two topographic positions using the GLM model restricted for temperature, rainfall and rainfall seasonality: (a) protected gullies, (b) ridges.
skewed nature of the response with a very sharp fall from the maximum to the extreme cannot be represented easily by either the GAM or GLM models. Discrepancies in temperature response between the models are less than for *E. cypellocarp*a. The smoothing function used in the GAM may need to be modified to include a smaller neighbourhood of observations. The GAM results for the nine eucalypt species (Austin and Meyers, unpublished data) support the skewed nature of the realised response to temperature concluded from the previous β-function study (Austin et al., 1994a).

The β-function used with GLM to model temperature response is constrained to zero at the limits and this also constrains the range of possible shapes between the limits. It is not appropriate to fit a β-function to rainfall and radiation because the upper and lower limits cannot be defined. Therefore log and polynomial functions were fitted in the GLM models (Austin et al., 1994a), but these are not necessarily better than a β-function.

The results for the different models for rainfall and radiation are less consistent than for temperature. The GLM polynomial or log functions often do not accord with the GAM model responses, though the GAM responses are generally consistent (Fig. 4(b) Fig. 7; Austin et al., 1994a) except close to the limits.

The practical management relevance is that use of unrestricted data may lead to the prediction of species where they do not occur. For conservation evaluation, this may constitute a conservative estimate. It may be better to predict presence when it is absent than to predict absence when it is present.

5. Discussion and conclusions

The case study presented here demonstrates some of the compromises often made when spatial modelling is done. The collated data set displays some spatial bias and some minor environmental combinations remain unsampled (Austin and Cawsey, unpublished data). However, the large number of plots sampled means that responses are well defined and numerous predictors can be used without running out of degrees of freedom. Lack of a suitable DTM, GIS and climatic estimates has meant that the full range of environmental process models of the type described by Moore et al. (1993) could not be used. However, the role of ecological theory and modern statistical regression models could be addressed.

Temperature has a direct impact on growth. Austin and Gaywood (1994) found that eucalypt species...
with predicted maxima occurring below 11.5°C had tails towards 11.5°C, whilst those with maxima above 11.5°C had tails towards lower temperatures. Twenty-one species out of a total of 24 which could be examined fitted this pattern.

The arbitrary and diverse response curves found for rainfall and radiation may reflect the fact that neither variable can be regarded as having a direct, simple impact on the distribution of species. While having predictive value, rainfall and radiation are more distal variables; they have less direct physiological effect than temperature, and fail to show consistent patterns between species. The response functions are subject to various distortions and biases due to the lack of appropriate environmental process models to link the variables into suitable scalars. Austin (1979) and Austin and Cunningham (1981) have previously demonstrated for *E. rossii* and *Eucalyptus mannifera* that better predictive models can be achieved by using a simple water-balance model to derive a moisture stress index rather than using mean annual rainfall or soil depth. Mean annual rainfall and radiation are not adequate surrogates for the proximal variables of moisture stress and light. Further progress is likely to be achieved by incorporating water balance estimates into the predictor variables (e.g. Moore et al., 1993).

The comparison of GAM and GLM models demonstrates that GAM has advantages over GLM due to the flexible nature of the non-parametric smoothing function. The influence of ‘naughty noughts’, however, can distort the response function and investigation of alternative model formulations such as restricting the data is necessary.

Niche theory postulates that complex curves may occur. Application of Occam’s razor suggests that consistent patterns between species due to common patterns of causation are more likely than each species responding in totally different ways to the same environmental predictor. Until more data sets with better GIS and environmental predictors have been analysed, it is premature to make too many claims about the realised niche response of species (cf. Austin et al., 1994b). It is clear, however, that generalisations based on better ecological theory of community composition along environmental gradients could greatly simplify the development of predictive models.

Robust predictive models are necessary for good management. Numerous non-statistical methods of predicting species distributions are currently being developed: for example, neural nets, genetic algorithm and decision trees (Austin et al., 1994b). These are likely to be sensitive to sample structure and the availability of suitable predictors. They also require experienced operators to achieve robust results. The nature of the predictions obtained by these methods needs detailed investigation (Austin, 1994). These computer induction methods are unlikely to replace statistical regression models where the data are suitable for statistical modelling. Differences in the shape of response curves, particularly at the limits of a species’ distribution can have major effects on the prediction of occurrence (Fig. 4(a) Fig. 6). Conservation evaluation often requires the comparison of unsampled areas, and so the assumption of particular shapes of response curve could lead to significant errors in the estimation of the conservation value of these areas.

Most aspects of forest management require adequate information on the distribution, abundance and growth potential of species. Statistical models of distribution utilising survey data, GIS and environmental predictors offer a powerful means of supplying that information in a form that can be subjected to rigorous testing. Development of models for species of greater perceived conservation value than eucalypts is occurring (Norton and Mitchell, 1993; Cork and Catling, 1996,) but there will be a constant need to examine the assumptions of these methods. Each of the requirements discussed here also needs constant re-assessment if cost-effective results are to be obtained.

Acknowledgements

We thank S.R. Morton, A.O. Nicholls, S.J. Cork and M. Barson for commenting on the manuscript. Part of this research was funded by the Australian Nature Conservation Agency.

References

Adomeit, E.M., Austin, M.P., Hutchinson, M.F. and Stein, J.L., 1987. Annual mean rainfall and temperature surfaces, and
contour maps of south-eastern Australia. CSIRO Institute of Natural Resources and Environment, Division of Water Resources Research, Canberra, Technical Memorandum 87/15.


Hastie, T., 1984. A summary of some surface fitting and contouring programs for noisy data. Consulting Report No. ACT 84/6, CSIRO Division of Mathematics and Statistics and Division of Water and Land Resources. CSIRO, Canberra, Australia.


