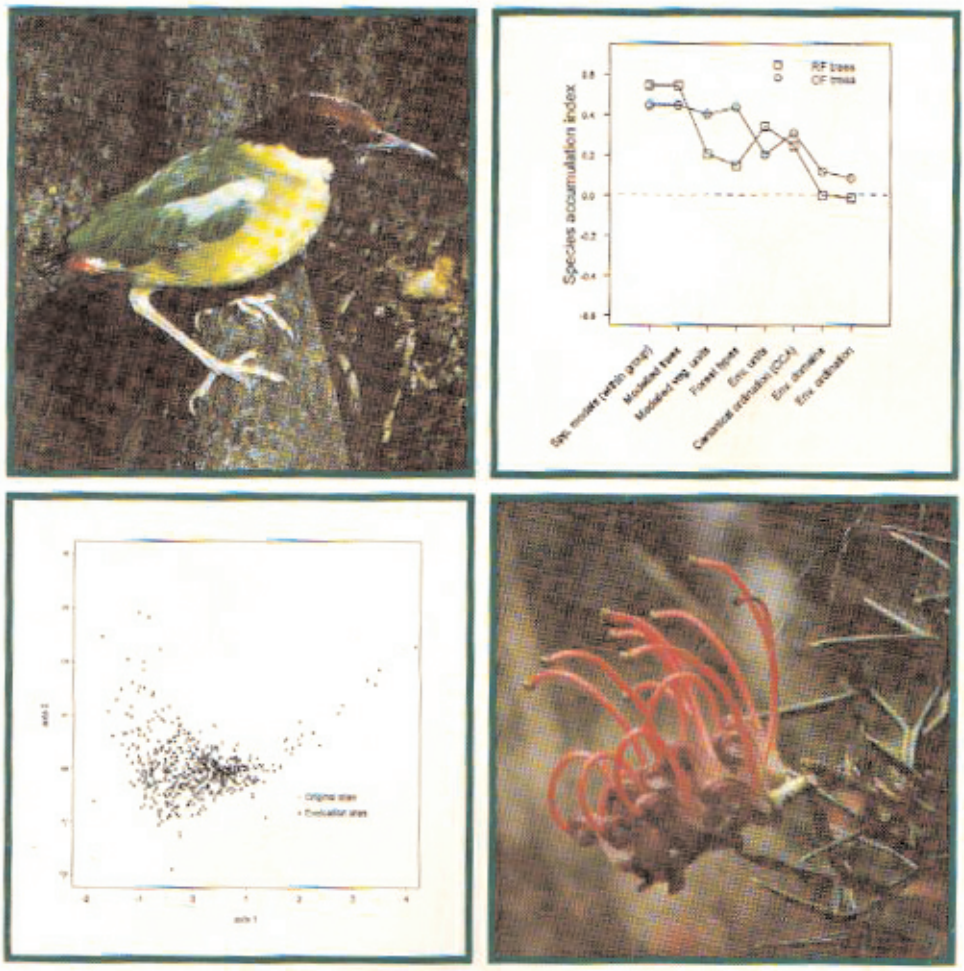
An Evaluation of the Effectiveness of Environmental Surrogates and Modelling Techniques in Predicting the Distribution of Biological Diversity



Consultancy report prepared by
Simon Ferrier and Graham Watson
NSW National Parks and Wildlife Service
for
Department of Environment, Sport and Territories

AN EVALUATION OF THE EFFECTIVENESS OF ENVIRONMENTAL SURROGATES AND MODELLING TECHNIQUES IN PREDICTING THE DISTRIBUTION OF BIOLOGICAL DIVERSITY

Consultancy report to the Biodiversity Convention and Strategy Section of the Biodiversity Group, Environment Australia

Simon Ferrier and Graham Watson

NSW National Parks and Wildlife Service

Authors:
Simon Ferrier and Graham Watson
NSW National Parks and Wildlife Service
Armidale District, PO Box 402, Armidale, NSW 2350

Front photographs:
Neville Fenton
NSW National Parks and Wildlife Service

ISBN 0 642 271283

Published by Environment Australia

Enquires, telephone 06-2500717 or 2500713
Biodiversity Group, Environment Australia

Cover design: NSW National Parks and Wildlife Service
Produced by: Australian Government Printing Service

1. Summary

The Department of Environment, Sport and Territories (DEST) contracted the NSW National Parks and Wildlife Service (NSW NPWS) to evaluate the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity.

The consultancy had two main components:

- evaluation of broad types of land mapping, environmental classification and ordination, and predictive modelling as general surrogates for biodiversity; and
- testing of the accuracy of different types of modelling in predicting species distributions.

For each of these components the consultant developed new analytical evaluation techniques and then applied these techniques to environmental and biological data from forested north east NSW and, to a lesser extent, arid north west NSW.

Two analytical techniques were developed for evaluating the general performance of surrogates, one based on matrix correlation and the other on species accumulation curves. Both approaches employ surrogate data and real biological survey data from a common set of field survey sites.

The matrix correlation approach assesses the correlation between two sites-by-sites distance matrices derived from the survey sites. One matrix measures the distance between sites in terms of the environmental surrogate under evaluation. The second matrix measures the distance between sites in terms of biological data collected by direct field survey.

The species accumulation approach simulates the building of a reserve system by selecting survey sites in a sequence that maximises representation of variation within the surrogate under consideration. The performance of the surrogate is assessed by using real biological data from the sites to calculate the cumulative number of species reserved after selecting each site. The results of this assessment are plotted as a species accumulation curve. This 'surrogate curve' is then used to derive an index of surrogate efficiency by relating the curve to two other reference curves: an 'optimum curve' derived by using the real biological data to select sites in a sequence that maximises the cumulative number of species reserved at each step, and a 'mean random curve' derived by selecting sites in random order , without reference to either the surrogate or the biological data.

The surrogate evaluation techniques were applied extensively to environmental and biological data from forested north east NSW. Surrogates evaluated included different types of vegetation mapping, abiotic environmental classification (domains), abiotic environmental ordination, canonical ordination and modelling of species distributions. These surrogates were assessed using field survey data for vascular plants (canopy and understorey), vertebrate animals (small reptiles, diurnal birds and microchiropteran bats) and invertebrate animals (ground-dwelling spiders, beetles and ants).

There was surprisingly little correlation between results obtained using the two indices of surrogate performance. The species accumulation approach was adopted as the basis for all subsequent analyses as it provided a more direct, and more readily interpretable measure of the performance of surrogates as a basis for reserve planning than did matrix correlation analysis.

The evaluation of surrogates in north east NSW revealed clear differences in performance, both between different types of surrogates and between different biological groups. All surrogates performed poorly for ground dwelling invertebrates. Reasonably consistent

differences in performance between broad types of surrogates were apparent for all vertebrate and vascular plant groups. The poorest performing surrogates were those derived purely from abiotic environmental data, i.e. environmental domain classifications and environmental ordinations. Canonical ordination (CCA) performed marginally better than purely abiotic classification and ordination for flora, but provided little improvement for vertebrate fauna. Forest type mapping generally outperformed all types of environmental classification and ordination for both vertebrate fauna and vascular plants.

The best overall performance of any surrogate in north east NSW was achieved by modelling of species distributions, especially when modelling was applied to species within the biological group under evaluation. The use of models for one biological group as a surrogate for other biological groups also appeared to have potential. In particular, modelled canopy trees performed better than forest type mapping as a surrogate for both canopy and understorey flora, and for vertebrates. Caution should be exercised in generalising these results from north east NSW to other regions, due to the potential effect of environmental variation and differences in data quality and availability.

A further trial application of the surrogate evaluation techniques to data from arid north west NSW demonstrated the applicability of these techniques to different environments and different types of surrogates. All mapped descriptors for the region (geology, landform, land systems, land units) performed very poorly as surrogates for vascular plants. This result highlighted the importance of considering regional differences when evaluating the performance of surrogates.

Specific analytical techniques were developed to evaluate the predictive accuracy of models of species distributions. These techniques can be used to assess the overall accuracy of models derived from presence/absence data and the discriminatory ability of models derived from either presence/absence or presence-only data. The techniques require independence between data used to fit a model and data used to evaluate that model. This independence is achieved either by applying jackkifing to the original dataset or by employing a truly independent test dataset.

The accuracy of predictive species models was evaluated using biological and environmental data from forested north east NSW. Fifty six species were used in the evaluation. These were drawn from a wide range of vertebrate and vascular plant groups and were selected to represent variation in frequency of occurrence (rare versus common) and strength of relationship with environmental variables used in the modelling. Models were fitted using both presence/absence and presence-only data, and both fine and coarse scale environmental data. Modelling algorithms evaluated included generalised linear modelling, generalised additive modelling, decision tree modelling and simple profile matching (BIOCLIM).

The evaluation of models in north east NSW revealed clear differences in performance between different modelling approaches. Models derived from presence/absence data performed better than models derived from presence-only data. Generalised additive modelling provided better performance than generalised linear modelling, which in turn performed better than decision tree modelling and BIOCLIM. Models performed better when derived using finer scale environmental data, especially in conjunction with accurately georeferenced presence/absence data. Modelling of species with strong environmental relationships performed better than modelling of species with weak environmental relationships.

The consultancy made the following recommendations (see Section 8.2 for detail):

- The evaluation techniques developed in this study should be employed as an interim standard for evaluating the performance of environmental surrogates and predictive modelling techniques.
- Further investigation of alternative approaches to the evaluation of environmental surrogates and predictive modelling techniques should be encouraged. Refinement

and extension of the techniques developed in this consultancy should also be facilitated.

- Results from the evaluation of surrogates and modelling techniques in forested north east NSW should be used as an interim guide to the relative performance of surrogates in similar forested environments. Caution should, however, be exercised in extrapolating these results to other regions, due to the potential effect of environmental variation and differences in data quality and availability.

- Evaluation of the performance of surrogates and modelling techniques elsewhere in Australia should be encouraged. The scope of such work should ideally cover a wide range of environments, spatial scales, biological groups and types of surrogates.

- Further research should be conducted into ways of refining surrogates to improve their performance as a basis for reserve planning.