

3. Evaluation of environmental surrogates: I. Development of evaluation methodology

3.1 Introduction

3.1.1 The role of surrogates in biodiversity conservation planning

Planning of conservation reserves requires information on species distribution and abundance. This information is also needed for environmental impact assessment and for monitoring the effects of environmental change. Direct surveys of species distribution and abundance are, however, expensive and time consuming. Regional biological surveys typically survey a very small proportion of a region's total area, usually sampling only a small number of scattered survey sites. Our knowledge of species distributions based on such surveys is therefore patchy and incomplete, both in terms of geographical coverage and in terms of taxonomic coverage (not all elements of biodiversity are surveyed at each site).

Geographical and taxonomic gaps in our direct knowledge of species distributions pose a major problem for the planning of conservation reserves. This problem is often addressed through the use of surrogate measures of species distributions (see Resource Assessment Commission 1993 for a recent review of surrogate use). Surrogate measures provide a means of dealing with information gaps in two different dimensions, taxonomic and geographical. A surrogate can address either the problem of taxonomic gaps, or the problem of geographical gaps, or both.

This consultancy report deals only with surrogates that address the problem of information gaps in the geographical dimension. Such surrogates provide complete geographical coverage of physical or biological information across a region and, assuming a correlation with real species distributions, serve as a convenient basis for reserve planning. The report evaluates surrogates in terms of their potential contribution to the selection of conservation reserves. To contribute effectively to reserve planning a surrogate must provide reasonably complete geographical coverage of large regions. Surrogates that rely on direct biological field survey of biodiversity, without any means of extrapolating these results across unsurveyed areas, are of limited value in reserve selection (but can be of great value in other fields of land management and planning). This report therefore does not evaluate the use of direct field survey of one component of biodiversity as a surrogate for direct survey of other components (e.g. field survey of flora as a surrogate for survey of fauna, or field survey of vertebrates as a surrogate for survey of invertebrates). Evaluation of this type of surrogacy is now starting to receive research attention elsewhere (e.g. Emberson 1985, Yen 1987, Landres *et al.* 1988, Kremen 1992, Oliver and Beattie 1993, 1996, Pearson 1994).

Many different types of surrogates have, or could, be used as a basis for conservation reserve planning. The most commonly used types can be classified into three broad groups:

- Mapped land classes derived from thematic mapping of a single attribute such as vegetation (e.g. Kuchler 1967) or soil (e.g. Isbell and McDonald 1984) or from intuitive integration of a suite of land attributes into a single map (e.g. land systems mapping, Christian and Stewart 1953; Mabbutt 1968; or mapping of bioregions, Thackway and Cresswell 1995).

- Environmental classifications or ordinations derived from numerical pattern analysis of primary environmental attributes (terrain, climate, substrate etc). The best known environmental classification approach used in Australia is that of environmental domain analysis (e.g. Mackey *et al.* 1988, 1989; Richards *et al.* 1990; Kirkpatrick and Brown 1991; Lewis *et al.* 1991; Thackway and Cresswell 1992; Belbin 1993a). The use of environmental ordination as a surrogate for biodiversity conservation has also been recently promoted in Australia by Faith and Walker (1993, 1996).
- Predictive models of the distribution of species, or potential habitat for species, derived by modelling biological survey data in relation to mapped environmental attributes (e.g. Austin *et al.* 1984; Stockwell *et al.* 1990; Ferrier 1991; Buckland and Elston 1993).

3.1.2 Evaluation of surrogate efficiency

As indicated above, the use of a surrogate as a basis for conservation reserve planning assumes a correlation between the surrogate and actual species distributions. It is further assumed that by designing a reserve system to represent variation within a surrogate, this reserve system will also represent individual species.

These assumptions have rarely been tested. Previous evaluations of surrogates have tended to focus on the correlation or association between biological survey data and a single type of surrogate, e.g. vegetation mapping (e.g. Braithwaite *et al.* 1988; Woinarski *et al.* 1988), landscape units (e.g. Greenslade and Thompson 1978), environmental domains (e.g. Kirkpatrick and Brown 1994), predictive species models (e.g. Lindenmayer *et al.* 1994). The purpose of most of these studies has been to demonstrate, and in some cases test the statistical significance of, correlation between the biota and the surrogate of interest (see also Pressey 1994 for a recent review of studies). Such testing contributes little, however, to our understanding of the effectiveness or efficiency of surrogates as a basis for reserve planning (although the pioneering work by Kirkpatrick and Brown 1994 is a notable exception). To evaluate the effectiveness of a surrogate we need not only to demonstrate a statistically significant correlation with the biota but also to measure the strength of that correlation. Furthermore we need to be able to compare the effectiveness of different types of surrogates.

This consultancy has developed a methodology for measuring the efficiency of a surrogate as a basis for reserve planning. The methodology can be applied to all major types of surrogate, thereby allowing evaluation of relative efficiency. The consultancy has applied the methodology to a wide range of surrogates using biological survey data from forested north-east NSW and arid north-west NSW (see Sections 4 and 5 of this report). The methodology is sufficiently generic to be employed in other regions.

3.2 Selection of evaluation methodology

3.2.1 Requirements

The evaluation methodology employed in this consultancy needed to satisfy a number of requirements. The first of these is a capability to evaluate surrogate efficiency by using species data collected at a sample of biological survey sites. It is assumed that the surrogates to be evaluated have also been mapped or estimated at each of these sites. It is further assumed that the sites are reasonably well spread across the variation within each surrogate, but not that every class within a surrogate contains sites. The biological data employed in this consultancy record presence or absence of species at each site. The analytical techniques reviewed and described below are, however, generally capable of accepting either presence/absence or species abundance data.

The evaluation methodology must analyse the relationship between two sets of data from a sample of survey sites: 1) the presence/absence of species at each site, and 2) the

mapped or estimated values of surrogates at each site. The complexity of the problem is greatly increased by the need to accommodate three very different types of surrogate data in the analyses:

- Surrogates for which each site is assigned one of a number of classes (e.g. vegetation types, environmental domains).
- Surrogates for which each site is assigned a position within a multidimensional surrogate space (e.g. environmental ordinations), possibly represented as a sites-by-sites similarity or dissimilarity matrix.
- Surrogates for which each site is assigned a number of predicted probabilities of species occurrence, one for each species modelled in a region.

To enable the efficiency of different surrogates to be compared, it is essential that the adopted methodology cater for all three data types.

A final requirement is that the methodology should produce an index of surrogate efficiency that indicates the likely performance of the surrogate as a basis for reserve selection, and that techniques are available for testing the statistical significance of this index and for estimating confidence limits.

3.2.2 Review of potential approaches

In selecting an analytical methodology for evaluating surrogate efficiency the consultancy reviewed a number of approaches from a wide range of disciplines:

- Techniques for assessing the predictive accuracy of models. Many analytical techniques have been developed for measuring the performance of predictive models using an independent validation dataset (e.g. Turner *et al.* 1989; Miller *et al.* 1991; Flather and King 1992; Power 1993). While these techniques can be used to evaluate the effectiveness of surrogates involving predictive modelling of species distributions, they are not applicable to surrogates for which no explicit prediction of species occurrence is being made (e.g. land systems, environmental domains). These evaluation techniques have been applied in a separate component of the consultancy, specifically evaluating the performance of different predictive modelling procedures (see Sections 6 and 7 of this report). They are not, however, suitable as a general methodology for comparing different types of surrogates.
- Techniques for assessing the agreement (or concordance) between two or more classifications of a common set of objects (e.g. Rohlf 1974, 1982; Faith 1986). While such techniques have some potential in assessing surrogate efficiency they require that each of the biological and surrogate datasets first be reduced to a classification, therefore resulting in loss of information. Of particular concern is the requirement that surrogates that are not already expressed as a classification must be reduced to that form. This will result in a considerable loss of information for surrogates such as environmental ordinations and modelled species distributions, and is therefore likely to penalise these surrogates in any comparative evaluation.
- Techniques for assessing the homogeneity of ground surveyed attributes within land classes and, conversely, the heterogeneity of these attributes between classes. This assessment can be performed using either simple univariate statistical procedures to analyse single species distributions (e.g. Braithwaite *et al.* 1988; Woinarski *et al.* 1988) or more sophisticated multivariate tools, such as homogeneity analysis (e.g. Pressey and Bedward 1991; Bedward *et al.* 1992), to analyse multiple species patterns. The approach has considerable potential for evaluating alternative land classifications but, like the previous approach, is not well suited for evaluating surrogates other than classifications (e.g. environmental ordinations, modelled species distributions).

- Techniques for assessing the correlation between two distance (or dissimilarity) matrices. Although matrix correlation analysis is a commonly used approach in other disciplines (e.g. Mantel 1967; Douglas and Endler 1982; Deitz 1983; Hubert 1985; Smouse *et al.* 1986) it has rarely been applied in ecology (although good examples are provided by Burgman 1987; Legendre and Fortin 1989; Leduc *et al.* 1992; Clark and Ainsworth 1993; Legendre 1993). The technique does, however, bear some similarity to other approaches more frequently applied in ecology such as canonical correlation (or correspondence) analysis, multidimensional similarity structure (or multidimensional scaling) analysis and Procrustes analysis. Matrix correlation analysis is well suited to the evaluation of surrogate efficiency because, unlike the other techniques listed above, it can accommodate very different types of surrogates. The only requirement is that the biological and surrogate data can each be expressed in the form of a matrix of distances between sites. In the case of the biological data this can be achieved using any of the numerous dissimilarity measures developed for this purpose. In the case of the surrogate data a distance matrix can be derived for all major types of surrogates without any substantial loss of information (details are provided in Section 3.3.2 below).
- Techniques for directly assessing the efficiency of a surrogate by simulating the formation of a reserve system based on that surrogate. All of the approaches described above assess the potential efficiency of a surrogate as a basis for reserve planning by measuring the correlation, or association, between that surrogate and real biological data. An alternative approach is to measure the performance of a surrogate as a basis for reserve planning directly. This can be achieved by using the surrogate data for the biological survey sites to build a simulated reserve system (i.e. selecting reserved sites in a sequence that maximises representation of variation within the surrogate) and then using the real biological data from these sites to assess the performance of the surrogate in terms of reservation of species. A simple technique for assessing this performance is to plot a species accumulation curve depicting the cumulative number of species reserved after the addition of each site to the simulated reserve system. Species accumulation curves (or 'species-area' curves) have been used widely in ecology for many different purposes (e.g. Palmer 1990; Colwell and Coddington 1994), including evaluation of different reserve selection strategies (e.g. Saetersdal *et al.* 1993), but do not appear to have been previously used to assess the performance of surrogates for biodiversity conservation.

3.2.3 Approaches adopted in DEST consultancy

We have used two approaches to evaluate the efficiency of surrogates, one based on matrix correlation and the other on the analysis of species accumulation curves. Matrix correlation analysis was selected as the most appropriate of several candidate techniques (listed above) for assessing the correlation, or association, between a surrogate and real biological data. Matrix correlation accommodates, and therefore facilitates comparison between, a wider range of surrogates than alternatives such as homogeneity analysis or analysis of agreement between classifications.

The analysis of species accumulation curves was selected as a technique for directly assessing the efficiency of surrogates in reserve planning. We have applied both the matrix correlation and species accumulation techniques to a wide range of biological datasets and surrogates, thereby enabling comparison of the two approaches (see Sections 4 and 5 of this report).

The two evaluation approaches used in the consultancy are now described in greater detail.

3.3 Approach 1: Matrix correlation analysis

3.3.1 Introduction to matrix correlation analysis

Mantel (1967) proposed a statistic for evaluating the correlation between two distance matrices, each based on the same set of objects. The significance of Mantel's statistic is tested using a generalised permutation, or Monte Carlo, strategy to overcome the problem of individual values in a distance matrix not being independent of one another. Since Mantel's original paper the technique has undergone considerable development and refinement. Non-parametric matrix correlation statistics have been introduced (Mantel and Valand 1970; Deitz 1983; Hubert 1985), allowing relaxation of the linearity assumption inherent in the original technique. The technique has also been extended to accommodate analysis of cophenetic distance matrices derived from hierarchical classifications (Hubert and Baker 1977; Lapointe and Legendre 1992), and partial correlation analysis of three or more matrices (Smouse *et al.* 1986). Good examples of the use of matrix correlation in ecological analysis are Burgman (1987,1988), Legendre and Fortin (1989), Leduc *et al.* (1992), Clark and Ainsworth (1993) and Legendre (1993).

For this consultancy, we used matrix correlation analysis to evaluate the correlation between two distance matrices derived for a common set of biological survey sites. The first matrix contains biological distances between all pairs of sites, estimated from the real species data recorded at those sites. The second matrix contains surrogate distances between all pairs of sites derived from data for the surrogate being evaluated.

3.3.2 Calculation of matrix correlation index of surrogate efficiency

Derivation of biological distance matrix

We estimated the biological distance between survey sites using the presence/absence version of the Bray-Curtis measure (Bray and Curtis 1957; the presence/absence version is also attributed to Czekanowski 1913) as recommended by Faith *et al.* (1987) and Belbin (1992) and now widely used for this purpose. The Bray-Curtis distance (D) between two sites ranges from 0.0 to 1.0 and, when based on presence/absence data, is expressed as

$$D = 1 - \frac{2A}{2A + B + C}$$

where A = number of species in common

B = number of species unique to the first site and

C = number of species unique to the second site.

Figure 3.1 gives an example of the application of this distance measure to a sample of biological data. We have derived separate biological distance matrices for a number of 'groups' of species: canopy plants, understorey plants, ants, spiders, beetles, reptiles, birds and bats. This has enabled evaluation and comparison of the performance of surrogates in relation to these groups.

Derivation of surrogate distance matrix

(a) sites by species table

	sp 1	sp 2	sp 3	sp 4	sp 5
site1	0	1	0	1	0
site2	1	0	1	1	1
site3	0	0	0	1	1
site4	1	1	0	0	1
site5	0	1	0	0	0

(b) sites by sites biological distance matrix

	site1	site2	site3	site4	site5
site1	0.00	0.75	0.50	0.60	0.33
site2	0.75	0.00	0.33	0.43	1.00
site3	0.50	0.33	0.00	0.60	1.00
site4	0.60	0.43	0.60	0.00	0.50
site5	0.33	1.00	1.00	0.50	0.00

Figure 3.1 Example of conversion of a sites by species presence/absence table into a sites by sites biological distance matrix using the Bray-Curtis measure

(a)

	site 2	site 3	site 4	site 5
site 1	0	1	1	0
site 2		1	1	0
site 3			0	1
site 4				1

(b)

	site 2	site 3	site 4	site 5
site 1	0.41	0.75	0.54	0.21
site 2		0.56	0.16	0.49
site 3			0.71	0.54
site 4				0.62

Figure 3.2 Examples of surrogate distance matrices for (a) a surrogate consisting of simple classes (e.g. vegetation types) and (b) a surrogate in which sites are assigned a position in multidimensional space (e.g. environmental ordination). Only the upper (off-diagonal) half of each symmetric matrix is depicted.

Surrogate distance matrices can be derived in a number of different ways depending on the type of surrogate under consideration:

- For surrogates consisting of simple mapped land classes (e.g. vegetation types, land systems) the distance matrix contains only binary values, 0 or 1. If two sites fall within the same land class they are assigned a distance of zero (0), otherwise they are assigned a distance of one (1). This is a commonly adopted approach for incorporating categorical data into matrix correlation analyses (e.g. Burgman 1987; Legendre and Fortin 1989). An example of a 0/1 surrogate distance matrix is provided in Figure 2. Where land classes are hierarchically structured, the coding of surrogate distance can be extended to reflect this structure using an approach proposed by Hubert and Baker (1977). For example if ‘forest types’ are grouped into ‘forest leagues’, which are in turn grouped into ‘forest groups’ then sites in the same forest type are assigned a distance of zero (0), sites in the same forest league but different forest types are assigned a distance of one (1), sites in the same forest group but different forest leagues are assigned a distance of two (2), and sites in different forest groups are assigned a distance of three (3).
- For surrogates derived by numerical classification of environmental attributes (e.g. environmental domains) a distance matrix can be derived in one of two ways. The first approach ignores information on the relationship between classes resulting from the classification by deriving a simple 0/1 distance matrix. If two sites fall within the same class they are assigned a distance of zero (0), otherwise they are assigned a distance of one (1). This approach is essentially testing the effectiveness of a surrogate for reserve planning if relationships between classes are ignored. However, as suggested by Faith and Walker (1993), information on these relationships is generally provided by classification procedures and should, wherever possible, be incorporated into reserve planning. We, therefore, evaluated surrogates derived from numerical classification of environmental data using an additional approach. In this approach the values used in the distance matrix are ultrametric, or cophenetic, distances between classes derived from the numerical classification procedure. For a detailed explanation of the approach see Rohlf (1982) or Lapointe and Legendre (1992).
- For surrogates derived from an ordination of environmental data, the values in the distance matrix are simply distances between sites in the fitted ordination space. An example of such a matrix is provided in Figure 3.2.
- For surrogates derived by modelling individual species distributions in relation to environmental variables the surrogate distance matrix is derived by estimating the biological distance (e.g. using the Bray-Curtis measure) between sites based on predicted probabilities for a set of modelled species, instead of real biological data. To ensure independence, the species models must be fitted using sites other than those employed in the surrogate evaluation. We achieve this independence by randomly dividing the sites into two equally sized datasets. One dataset is used to fit models, and the other to evaluate the efficiency of these models as a surrogate. The Bray-Curtis distance (D) between sites is calculated from predicted species probabilities as

$$D = \frac{\sum_{k=1}^n |p_{ik} - p_{jk}|}{\sum_{k=1}^n (p_{ik} + p_{jk})}$$

where p_{ik} = predicted probability of species k at site i

p_{jk} = predicted probability of species k at site j

n = number of species

Calculation of matrix correlation coefficient

Mantel's (1967) original statistic for measuring matrix correlation was an unnormalised Pearson product-moment correlation coefficient. This statistic assumes a linear relationship between the two distance matrices under evaluation, an assumption unlikely to be met by most ecological data (Faith *et al.* 1987). This assumption can be relaxed through use of non-parametric correlation coefficients in matrix correlation analysis, as demonstrated by Mantel and Valand (1970), Deitz (1983) and Hubert (1985).

The statistic we used to measure correlation between surrogate and biological distance matrices is Spearman's non-parametric correlation statistic (Spearman 1904) designated by ρ (*rho*). Spearman's correlation is used in preference to the other commonly applied non-parametric correlation statistic, Kendall's *tau*, because the latter proved to be too computationally demanding for analysis of very large distance matrices. Deitz (1983) has conducted a detailed evaluation of the performance of Pearson's, Spearman's and Kendall's statistics in matrix correlation analysis and concluded that while Kendall's statistic has greater power than Spearman's the difference is only marginal (relative to the greatly improved performance of both non-parametric statistics over Pearson's statistic).

Calculation of Spearman's correlation statistic is based on a ranking of the distances within each of the two distance matrices. Figure 3.3 provides an example of the transition of distance matrices into the ranked form required for calculation of the correlation statistic. As recommended by Conover (1980), three alternative algorithms are used to calculate Spearman's statistic, the choice depending on the incidence of ties in the rankings of the two matrices. If there are no ties, ρ is calculated as

$$\rho = \frac{\sum_{i=1}^n \left[R(X_i) - \frac{n+1}{2} \right] \left[R(Y_i) - \frac{n+1}{2} \right]}{n(n^2 - 1) / 12}$$

if there are few ties (<20%)

$$\rho = 1 - \frac{6 \sum_{i=1}^n [R(X_i) - R(Y_i)]^2}{n(n^2 - 1)}$$

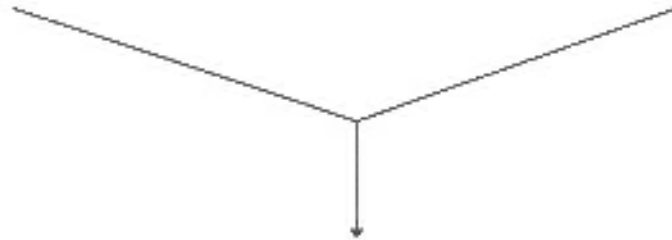
and if there are many ties (>20%)

surrogate distance matrix

	site 1	site 2	site 3	site 4	site 5
site 1	0		0.75	0.54	0.21
site 2	0.41	0	0.56	0.16	0.49
site 3	0.75	0.56	0		0.54
site 4	0.54	1.16	0.71	0	0.62
site 5	0.21	0.49	0.54	0.62	0

biological distance matrix

	site 1	site 2	site 3	site 4	site 5
site 1	0.0				
site 2	0.5	0.0			
site 3	1.0	1.0	0.0		
site 4	0.33	1.0	0.6	0.0	
site 5	1.0	1.0	0.33	0.5	0.0



Off-diagonal values

Surrogate distance	0.41	0.75	0.54	0.21	0.56	0.16	0.49	0.71	0.54	0.62
Biological distance	0.5	1.0	0.33	1.0	1.0	1.0	1.0	0.6	0.33	0.5



Rank transformations

Surrogate distance	3	10	5.5	2	7	1	4	9	5.5	8
Biological distance	3.5	8	1.5	8	8	8	8	5	1.5	3.5

Figure 3.3 An example of the transition from distance matrices to ranked data in preparation for calculation of the Spearman correlation coefficient. The off-diagonal values have been extracted in row sequence. Rank assignments commence at 1 and in the case of ties, the ranks assigned are the average of the rank positions covered by the ties.

$$\rho = \frac{\sum_{i=1}^n R(X_i)R(Y_i) - n\left(\frac{n+1}{2}\right)^2}{\left(\sum_{i=1}^n R(X_i)^2 - n\left(\frac{n+1}{2}\right)^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n R(Y_i)^2 - n\left(\frac{n+1}{2}\right)^2\right)^{\frac{1}{2}}}$$

which is simply Pearson's r computed on the ranks and average ranks.

The correlation between the surrogate and biological distance matrices can also be presented graphically in the form of a scatterplot in which each point represents a pair of survey sites plotted against two axes, one representing surrogate distance and the other biological distance. Examples are provided in Figures 3.4 and 3.5.

3.3.3 Significance test and confidence limits for matrix correlation index

Significance test

A significance test is required to test whether an obtained matrix correlation index of surrogate efficiency is significantly greater than zero. The significance of matrix correlation coefficients is usually tested using a generalised permutation, or Monte-Carlo, strategy to overcome the problem of individual values in a distance matrix not being independent of one another (Mantel 1967; Mielke 1978; Manly 1991). We test each coefficient against 1000 random permutations, as recommended by Manly (1991) and Lapointe and Legendre (1992). In each permutation the biological distance matrix is held constant while the rows of the surrogate distance matrix are permuted. The matrix correlation between the two matrices is calculated after each permutation.

For the determination of significance, a one-sided test is required as only positive correlations are of interest. The significance level of the observed correlation based on 1000 permutations is, therefore,

$$\frac{\left(\sum_{i=1}^{1000} I_i\right) + 1}{1000 + 1}$$

where $I_i = 1$ if $\rho_i \geq \rho_0$

$I_i = 0$ if $\rho_i < \rho_0$

ρ_i = matrix correlation after permutation i and

ρ_0 = matrix correlation using the observed data.

This significance level can be interpreted as the probability that a matrix correlation coefficient equal to, or greater than, the observed value could have been obtained by chance.

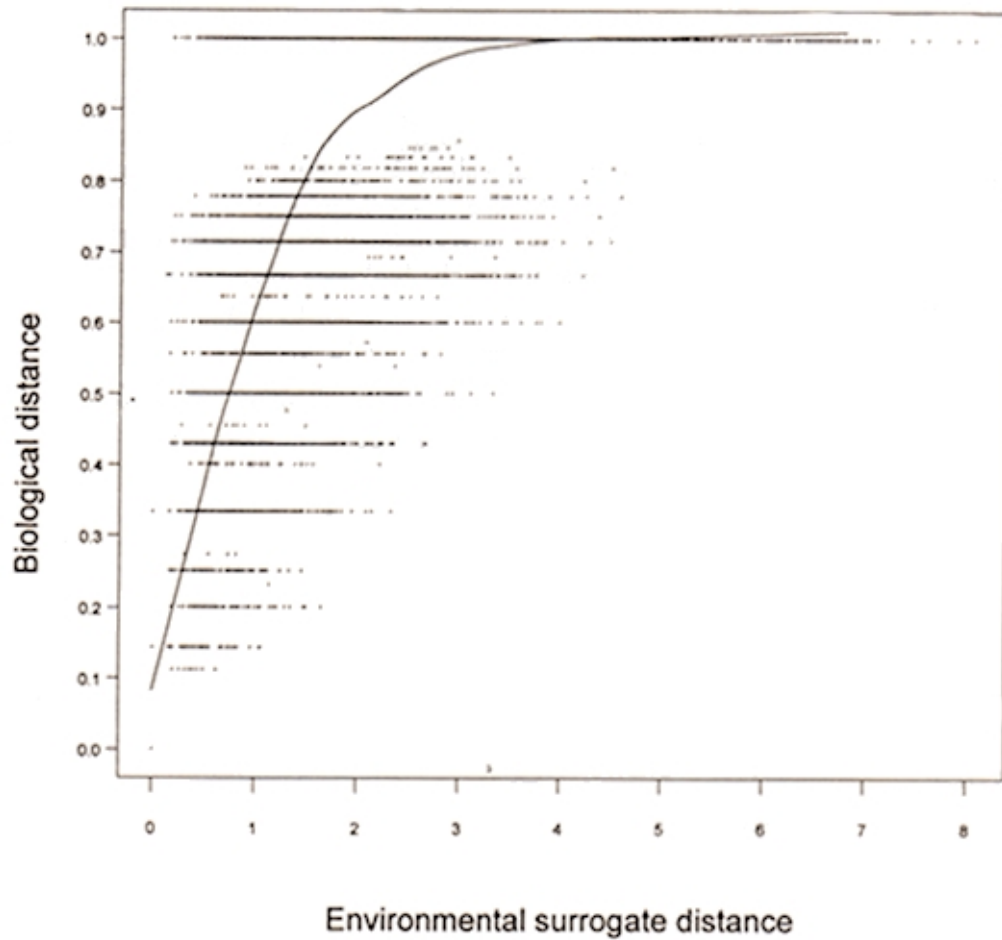


Figure 3.4 Example of a scatterplot depicting the correlation between surrogate distance and biological distance, for a surrogate yielding continuous distance values. Each dot represents a pair of survey sites. The fitted curve is a LOESS smoothed regressions line plotted for illustrative purpose only.

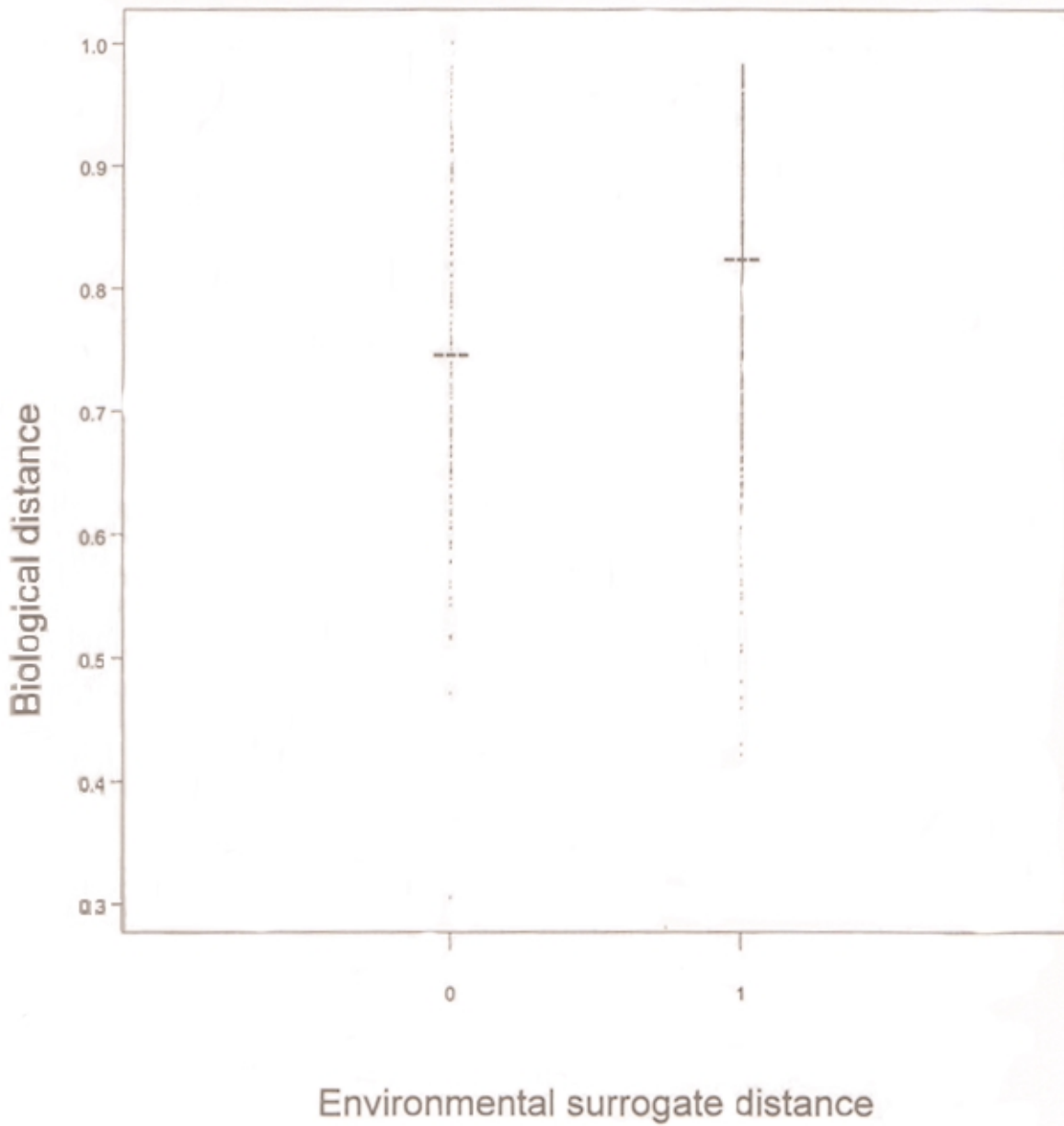


Figure 3.5 Example of a scatterplot depicting the correlation between surrogate distance and biological distance, for a surrogate consisting of classes (and therefore only 0 or 1 surrogate distances). Each dot represents a pair of survey sites. The horizontal line represent mean values.

Confidence limits

Confidence limits for each obtained matrix correlation index are estimated using bootstrapping (Efron 1979; Efron 1982; Efron and Tibshirani 1993). We use 100 bootstrap samples to estimate confidence limits (based on Efron and Tibshirani's 1993 recommendation that between 30 and 200 samples should produce a good approximation of a statistic). In each sample, sites are randomly selected with replacement until the number of sites in the sample equals the number of sites used in the original matrix correlation analysis. Each bootstrap sample is used to calculate a matrix correlation coefficient. These bootstrapped coefficients are then used to estimate a standard error and 95% confidence limits for the observed correlation coefficient (see Efron and Tibshirani 1993 for details).

3.4 Approach 2: Species accumulation analysis

3.4.1 Introduction to species accumulation analysis

Species accumulation curves have been used widely in ecology for many different purposes (e.g. Palmer 1990; Colwell and Coddington 1994), including evaluation of different reserve selection strategies (e.g. Saetersdal *et al.* 1993), but do not appear to have been previously used to assess the performance of surrogates for biodiversity conservation.

We use each surrogate to build a simulated reserve system by selecting survey sites in a sequence that maximises representation of variation within the surrogate. The performance of the surrogate is then assessed using the real biological data from the survey 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.

The species accumulation curve yielded by a surrogate is then used to derive an index of surrogate efficiency by relating this curve to two other 'reference' curves: 1) an 'optimum' species accumulation 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 2) a 'mean random' species accumulation curve representing the mean of a sample of species accumulation curves derived by selecting sites in random order, without reference to either the surrogate or the biological data. An example of the three species accumulation curves used to evaluate surrogate efficiency is provided in Figure 3.6.

3.4.2 Calculation of species accumulation index of surrogate efficiency

Derivation of surrogate curve

As indicated above, a 'surrogate curve' is derived by selecting survey sites in a sequence that maximises representation of variation within the surrogate. We used three different algorithms to select sites depending on the type of surrogate:

- For surrogates consisting of simple mapped land classes (e.g. vegetation types, land systems), and yielding a binary 0/1 distance matrix (see Section 3.3.2 for details), a very basic selection algorithm is employed. One site is randomly selected from each land class in turn, starting with the most common class (i.e. the class with the most sites) and working through in descending order to the rarest class. This procedure is repeated until no sites remain. The strategy at each step is therefore to select a site from the land class that will result in the most equitable division of reserved sites between land classes. If two or more classes can contribute equally at any given step then a site is chosen from the most common class. Because sites are selected at random from within each class the algorithm will produce a slightly different result each time it is applied. This problem is partly addressed by the use of bootstrapping

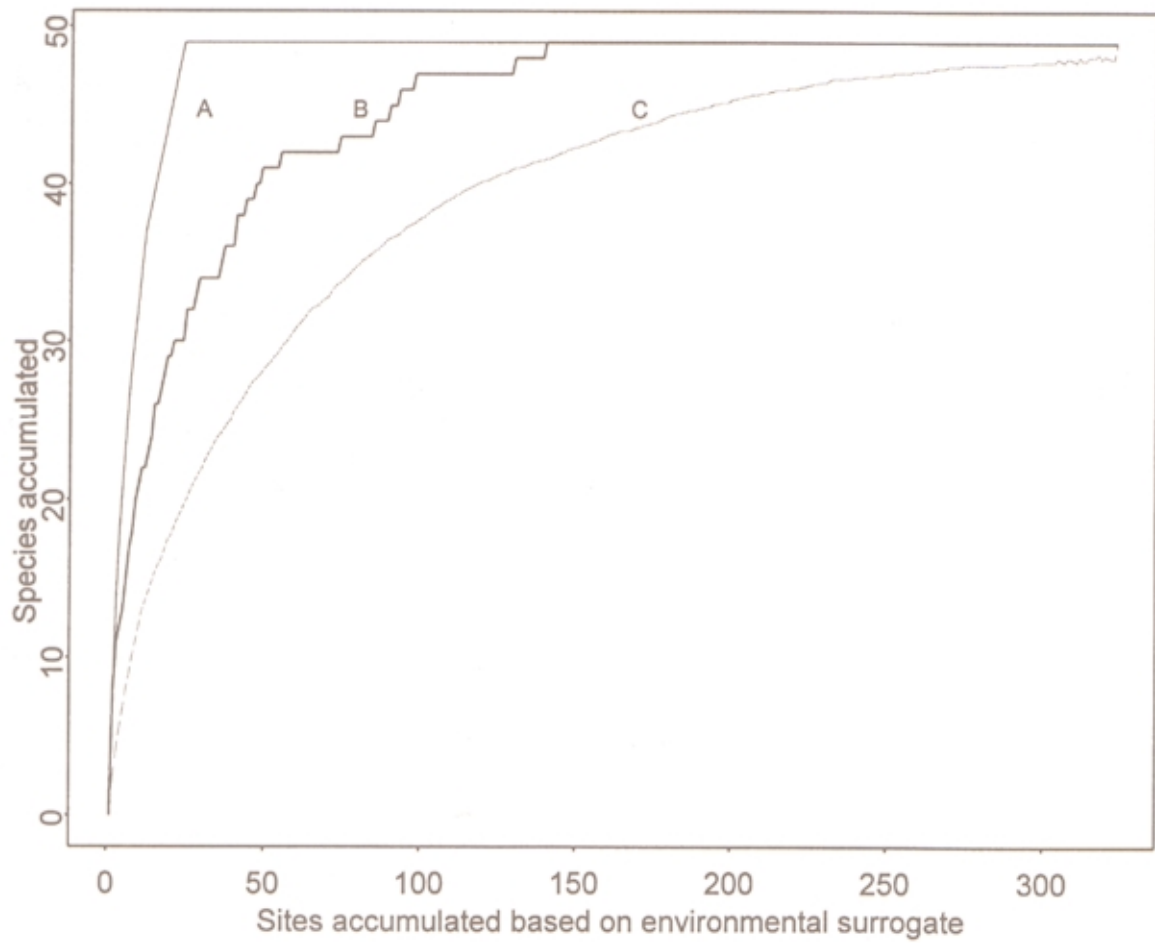


Figure 3.6 Example of the three species accumulation curves used to evaluate the efficiency of a surrogate. A = optimum reference curve, B = surrogate curve, C = mean random reference curve

to calculate confidence limits for the species accumulation index, as detailed below in Section 3.4.3.

- For surrogates yielding a non-binary distance matrix the discrete p-median selection algorithm is employed, as recommended by Faith and Walker (1993, 1996). Surrogates catered for by this approach include environmental ordinations, and classifications for which ultrametric, or cophenetic, distances between classes are available. The p-median algorithm is based on operations research principles relating to the problem of optimum location of objects (see Tansel *et al.* 1983). Faith and Walker (1993, 1996) have demonstrated that the discrete p-median approach has excellent properties in terms of capturing the diversity of an environmental surrogate. The strategy of the discrete p-median approach is to select, at each step, the site which minimises the average environmental (surrogate) distance from all unselected sites to their nearest selected site. Faith and Walker (1993, 1996) provide a comprehensive description of the procedure.
- For surrogates derived by modelling individual species distributions in relation to environmental variables two approaches can be used to select sites. One approach is to derive a surrogate distance matrix by estimating the biological distance (e.g. using the Bray-Curtis measure) between sites based on predicted probabilities for a set of modelled species, instead of real biological data. Sites can then be selected using the p-median algorithm just described. An alternative approach is to use the predicted species probabilities directly in the selection of sites. This approach is illustrated in Figure 3.7. The approach employs standard probability theory to estimate the probability that a species is reserved at least once in a set of sites given the predicted probability of occurrence at each site. The sum of the estimated probabilities of reservation for the individual species also provides an estimate of the total number of species reserved. The algorithm proceeds by, at each step, selecting the candidate site that makes the greatest contribution to the estimated total number of reserved species. This is achieved by first estimating, for each candidate site in turn, the probability $P(X)$ that each species will be reserved at least once if the reserve system is extended to include that site:

$$P(X) = P(X_{S_1}) + P(X_{S_2}) - P(X_{S_1} + X_{S_2})$$

where $P(X_{S_1})$ is the probability of the species occurring in the existing reserve system (before adding a new site)

$P(X_{S_2})$ is the probability of the species occurring at the candidate site

$P(X_{S_1} + X_{S_2})$ is calculated in the usual manner as the product of $P(X_{S_1})$ and $P(X_{S_2})$.

To ensure independence, the species models must be fitted using sites other than those employed in the surrogate evaluation. We achieve this independence by randomly dividing the sites into two equally sized datasets. One dataset is used to fit models, and the other to evaluate the efficiency of these models as a surrogate.

Derivation of optimum curve

The 'optimum curve' is derived by simply using the real biological data to select sites in a sequence that maximises the cumulative number of species reserved at each step. In other words, the first site selected is that with the greatest total number of species. The second

STEP 1: Selection of first site

	sp 1	sp 2	sp 3	sp 4	sp 5	total	
site1	0.4	0.1	0.3	0.2	0.2	1.2	
site2	0.3	0.4	0.4	0.3	0.4	1.8	
site3	0.1	0.2	0.5	0.5	0.1	1.4	
site4	0.2	0.4	0.8	0.3	0.3	2.0	site 4 selected
site5	0.4	0.3	0.4	0.2	0.3	1.6	
site6	0.1	0.2	0.3	0.4	0.1	1.1	
site7	0.1	0.3	0.4	0.2	0.1	1.1	

STEP 2: Selection of second site (site 4 already reserved)

	sp 1	sp 2	sp 3	sp 4	sp 5	total	
site1	0.52	0.46	0.86	0.44	0.44	2.72	
site2	0.44	0.64	0.88	0.51	0.58	3.05	site 2 selected
site3	0.28	0.52	0.9	0.65	0.27	2.62	
site5	0.52	0.58	0.88	0.44	0.51	2.93	
site6	0.28	0.52	0.86	0.58	0.27	2.51	
site7	0.28	0.58	0.88	0.44	0.27	2.45	

STEP 3: Selection of third site (sites 4 and 2 already reserved)

	sp 1	sp 2	sp 3	sp 4	sp 5	total	
site1	0.73	0.81	0.98	0.73	0.76	4.01	
site3	0.60	0.83	0.99	0.83	0.69	3.94	
site5	0.73	0.85	0.99	0.73	0.79	4.09	site 5 selected
site6	0.60	0.83	0.98	0.79	0.69	3.89	
site7	0.60	0.85	0.99	0.73	0.69	3.86	

Figure 3.7 Example application of algorithm for selecting sites based on predicted probabilities of species occurrence. The values in the site by species table for Step 1 are probabilities of occurrence as predicted by a model for each species. These probabilities are summed for each site, and the site with the highest sum is selected for reservation (site 4). The values in the table for Step 2 are cumulative probabilities of reservation for each species if one of the remaining sites is added to the reserve system, assuming that site 4 is already reserved. Step 3 is a repetition of Step 2 with two sites now reserved. The procedure is repeated until no sites remain. Details are provided in the text.

site selected is that with the greatest number of species not recorded at the first site, and so on.

This is a 'greedy' selection algorithm in that it never looks more than one step ahead when selecting sites. The resultant curve therefore represents the 'optimum' accumulation of species that can be achieved using a simple greedy selection algorithm. More sophisticated selection algorithms (e.g. integer programming, genetic algorithms) could be used to improve the performance of the accumulation curve, and to identify a globally optimum curve from all possible selection sequences. Such algorithms were too computationally demanding to be practically applied in this study (especially given the use of bootstrapping for calculation of confidence limits, described below). Furthermore, the use of a 'non-greedy' algorithm for deriving the optimum curve would not have provided a fair comparison with the surrogate curves, all of which are themselves based on greedy algorithms.

Derivation of mean random curve

The 'mean random curve' represents the mean of a sample of species accumulation curves derived by selecting sites in random order, without reference to either the surrogate or the biological data. We calculate the mean random curve from a sample of 1000 random curves. This sample is not only used to calculate the mean random curve but also to estimate the statistical significance of the species accumulation index (see Section 3.4.3 below).

Calculation of the species accumulation index

Once all three species accumulation curves have been derived, the efficiency of a surrogate is measured by calculating the 'species accumulation index' as:

$$(s - r) / (o - r)$$

where s = log area under the surrogate curve

r = log area under the mean random curve

o = log area under the optimum curve

The species accumulation index can range from one (1) for a 'perfect' surrogate down to zero (0) or less for a surrogate that performs no better than a random selection of sites. A number of other measures of surrogate efficiency could be derived from the species accumulation curves; for example, the number of sites required to reserve 90% of species. Further research on the performance of such alternatives is required.

3.4.3 Significance test and confidence limits for species accumulation index

Significance test

The sample of 1000 random curves used to calculate the mean random curve (see section 4.2 above) can also be used to perform a Monte-Carlo test (Manly 1991) of the statistical significance of the surrogate curve. This test assesses whether the area under the species accumulation curve for the surrogate is significantly greater than that likely to be obtained by selecting sites in random order.

The significance level of the surrogate curve in relation to 1000 random curves is

$$\frac{\left(\sum_{i=1}^{1000} I_i\right)+1}{1000+1}$$

where $I_i = 1$ if $a_i \geq a_0$
 $I_i = 0$ if $a_i < a_0$
 $a_i = \log$ area under random curve i and
 $a_0 = \log$ area under the surrogate curve.

This significance level can be interpreted as the probability that an area equal to, or greater than, that under the surrogate curve could have been obtained by chance.

The sample of 1000 random curves can also be used to provide a graphical indication of the significance of a surrogate curve as illustrated in Figure 3.8.

Confidence limits

Confidence limits for each species accumulation index obtained are estimated using bootstrapping (Efron 1979; Efron 1982; Efron and Tibshirani 1993). We use 100 bootstrap samples to estimate confidence limits (based on Efron and Tibshirani's 1993 recommendation that between 30 and 200 samples should produce a good approximation of a statistic). In each sample, sites are randomly selected, with replacement, until the number of sites in the sample equals the number of sites used in the original species accumulation analysis. Each bootstrap sample is used to derive a surrogate curve, optimum curve and mean random curve. The species accumulation indices calculated from these curves are then used to estimate a standard error and 95% confidence limits for the observed index (see Efron and Tibshirani 1993 for details).

3.5 Software

Both evaluation techniques (matrix correlation and species accumulation) were implemented using special-purpose software developed as part of this consultancy. The software was initially developed as a series of functions within the Microsoft Windows version of the S-PLUS statistical package (Statistical Sciences Inc. 1994). The speed of processing achieved using S-PLUS was not fast enough to enable the planned surrogate evaluations to be completed within the time period allocated to the consultancy. The software was therefore redeveloped as a collection of FORTRAN-77 programs using the Lahey compiler (Lahey Computer Systems Inc. 1992). The FORTRAN version of the software achieved acceptable processing speeds, especially when run on Pentium PCs. The S-PLUS version of the software has been retained as an alternative tool for calculating the indices of surrogate performance. This has proved particularly useful as a means of checking results obtained from the FORTRAN software.

Two input files are required by the matrix correlation software, one containing a biological distance matrix and the other containing a surrogate distance matrix (for the same set of survey sites). Two input files are also required by the species accumulation software. The first file is a sites-by-species matrix recording the presence or absence of individual species at the survey sites. The structure of the second file depends on the type of surrogate being evaluated. If the surrogate consists of simple land classes the file

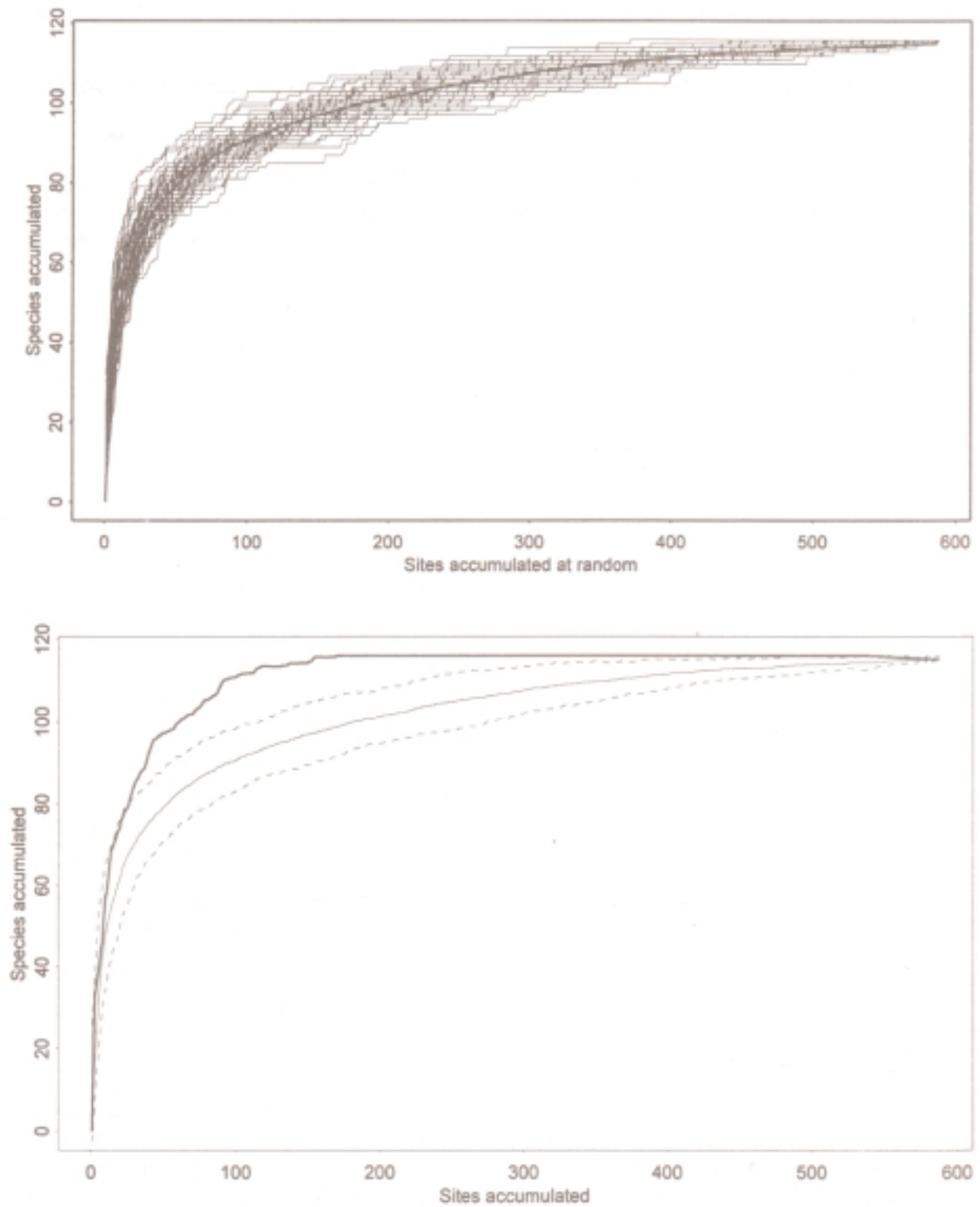


Figure 3.8 Example of the derivation of a mean random reference curve with 95% confidence limits (bottom graph), from a sample of random species accumulation curves (top graph). A surrogate curve is superimposed on the bottom graph (bold line).

contains a vector of class codes, one for each survey site. If the surrogate measures environmental distance between sites the file contains a surrogate distance matrix. For surrogates using probabilistic predictions from species habitat models the file is a sites-by-species matrix containing the predicted probabilities of species occurring at the survey sites.

3.6 Discussion

In this consultancy we have applied both measures of surrogate efficiency (matrix correlation and species accumulation analysis) to a wide range of surrogates and biological datasets. This has facilitated comparison of the two approaches. Preliminary results presented in the Progress Report (June 1995) and more extensive results presented in Section 4 of this report indicate that the two measures are only very weakly correlated. Surrogates performing well in matrix correlation analysis often perform badly in species accumulation analysis, and *vice versa*.

This lack of concordance between the two techniques is perhaps disappointing, but not necessarily surprising. The two techniques are evaluating surrogate efficiency in quite different ways. Matrix correlation analysis is evaluating the correlation between a surrogate and the biota in terms of a general measure of biological dissimilarity. This approach assumes that the general measure of biological dissimilarity used to evaluate surrogates is itself a good basis for reserve planning. Species accumulation analysis is instead directly evaluating the performance of a surrogate in reserve planning in terms of the efficiency with which species would be represented within a reserve system based on that surrogate. The lack of concordance between results obtained using the two techniques therefore raises an interesting question concerning the appropriateness of using general measures of biological dissimilarity (and classifications or ordinations derived from these) as a basis for reserve planning. A related issue concerns the choice of a specific dissimilarity measure. Would a measure other than the Bray-Curtis association measure provide a better basis for evaluating surrogate efficiency and better concordance between the two evaluation approaches? Further research is clearly needed to address these issues.

The issue of poor concordance between the two measures of surrogate performance is analysed and discussed in greater detail in Section 4.4.1. The conclusion reached in that section is that species accumulation analysis should be employed in preference to matrix correlation analysis as a measure of surrogate efficiency, at least in relation to reserve planning. Species accumulation analysis provides a direct, and readily interpretable, measure of the performance of surrogates.

There is considerable scope for refining, or extending, both of the techniques developed in this consultancy. Refinements worthy of further investigation include:

- Attributes other than species presence/absence could be used to evaluate the performance of surrogates as a basis for reserve planning. Such attributes include species abundance, genetic variation within species and community composition. These attributes could be used to evaluate the ability of surrogates to distinguish between sites with the same species present, but differing in the relative abundance (or dominance) of these species, and therefore community composition.
- The techniques used in this study have been designed to assess the performance of surrogates when applied individually to reserve planning. However, as demonstrated by Kirkpatrick and Brown (1994), reserve planning might be most effectively achieved by using two or more surrogates in combination (e.g. vegetation mapping and environmental domains). The developed techniques should be extended to evaluate the performance of such combinations.
- The matrix correlation approach could benefit from further investigation of alternative measures of biological distance, and alternative measures of correlation between the

two distance matrices. In particular, a better solution is needed for dealing with surrogate distance matrices containing only zeros and ones.

- The species accumulation approach could be modified to provide a more realistic simulation of reserve planning. Real reserve planning involves the selection of sizable blocks of land, not individual survey sites. To better simulate this process the evaluation methodology perhaps needs to select sites in groups (of neighbouring sites) rather than individually.
- More sophisticated ‘non-greedy’ algorithms need to be investigated for deriving the surrogate and optimum species accumulation curves.
- Other variables derived from the species presence/absence data could be used in place of the ‘number of species reserved’ variable as the vertical axis of species accumulation curves. Examples include the number of species for which reservation of a specified proportion of sites (with that species present) has been achieved, or the mean proportional reservation across all species.
- Other techniques could be used to measure the efficiency of species accumulation curves. For example, the number of sites required to reserve a specified proportion of species (e.g. 90%).
- Further research is needed to assess the extent to which the different selection algorithms used to derive species accumulation curves for different types of surrogates unfairly biases the comparison of these surrogates. Less biased selection techniques may need to be developed.
- Variations on the discrete p-median selection algorithm might be more appropriate for evaluating some types of surrogates. Candidate algorithms include the continuous p-median algorithm (Faith and Walker 1993, 1996) and the phylogenetic diversity criterion (Faith 1994).

As indicated earlier in this chapter there are many different ways of approaching the evaluation of surrogate efficiency. The techniques we have adopted are just two of many possible alternatives. They should not be regarded as the definitive solution to surrogate evaluation. Further research in this area is badly needed and should be encouraged.

The two techniques proposed nevertheless provide an interim basis for evaluating surrogate efficiency. This consultancy has already applied the techniques to a wide range of surrogates and biological dataset from forested north-east NSW and arid north-west NSW (see Sections 4 and 5). The tools are sufficiently generic to be applied to other surrogates in other regions. The only requirement is a dataset containing both surrogate data and real species data for the same set of biological survey sites.

Figure 4.1 North east NSW study area.

