Option Value, Gallot's Inequality, and the Measurement of Biological Diversity

Stephen Polasky
and
Andrew R. Solow
(Woods Hole Oceanographic Institution)

Working Paper No. 241
September 1992

OPTION VALUE, GALLOT'S INEQUALITY, AND THE MEASUREMENT OF BIOLOGICAL DIVERSITY

Andrew R. Solow

Woods Hole Oceanographic Institution

Stephen Polasky

Department of Economics

Boston College

ABSTRACT

To allocate scarce resources to the conservation of biological diversity in an effective way, it is necessary to measure diversity. One argument in favor of species conservation is that species have an option value arising from their potential future use (e.g., in curing a disease). We use a simple model that relates the genetic distance between species to a nonparametric lower bound on option value. This leads to a simple measure that can be used to compare different sets of species. The measure can be interpreted as the effective number of species in the set. An example is given illustrating the use of this measure in comparing different sets of glucosinolate-producing plants.

1. INTRODUCTION

One of the most striking features of life on Earth is its diversity. When human activities contribute to the extinction of plant or animal species, part of this diversity is lost. A number of national and international measures have already been taken to protect biological diversity and this issue promises to be among the foremost environmental issues of this decade and beyond.

As long as the resources devoted to the conservation of biological diversity are limited, it is necessary to allocate them across conservation projects. A small but growing body of work has focused on optimizing this allocation (e.g., Vane-Wright, et al., 1991; Weitzman, 1992; Solow, et al., 1992). This work essentially assumes that diversity per se has value and goes on to address the central issue of measuring diversity. In keeping with this literature, for the purposes of this paper the diversity of a set of species refers in qualitative terms to the joint dissimilarity of the species within the set. It is important to note that this usage is different from that in ecology (e.g., Pielou, 1975), in which diversity is viewed as a property of the abundance distribution. The information available for constructing a measure of diversity consists of the pairwise genetic distances between the species.

It is commonly argued that diversity should be conserved not only for its own sake but for utilitarian reasons. For example, species have an option value arising from their potential future

use in resisting a pest or curing a disease (e.g., Brown and Goldstein, 1984). Polasky, et al. (1992) used a simple search model to show that such considerations do impart a value to diversity, provided that similar species are more likely to be substitutes than dissimilar species. Similarly, Weitzman (1992) showed that his measure of diversity is linked to a lower bound for option value under a simple evolutionary model. In each case, the linkage between diversity and option value depends on a rather specific model. The purpose of this paper is to make this linkage in a nonparametric setting.

The remainder of this paper is organized as follows. In Section 2, we use a nonparametric lower bound for option value to derive a measure by which different sets of species can be compared. In Section 3, we discuss the relationship between this measure and diversity. A simple example is given in Section 4. Section 5 contains some concluding remarks.

2. A MEASURE BASED ON OPTION VALUE

Suppose that we are interested in conserving species because of the possibility that they will provide a specific benefit in the future. The essential property of this benefit is that having more than one species that provide it is no better than having a single species that provides it. To a first approximation, an example of such a benefit is a cure for cancer and, for concreteness, we will use this as a metaphor.

Consider a set of species $S = (s_1, s_2, \ldots, s_n)$ and let B_i be the event that s_i is a cure. The event that S contains a cure is $B(S) = \bigcup_S B_i$ and the option value of S is R p(S) where R is the value of a cure and p(S) = prob(B(S)). This probability provides a basis for comparing S to other sets of species.

In the absence of specific information, it is reasonable to assume that $prob(B_i) = p$, $i = 1, 2, \ldots, n$. We will make the further assumption that:

$$prob(B_i \mid B_j) = p + (1-p)\, f(d_{ij}) \qquad (1)$$

where d_{ij} is the genetic distance between s_i and s_j and f(d) satisfies the following conditions:

$$f(0) = 1$$

$$f(\infty) = 0$$

$$f' \leq 0$$

To ensure consistency, we will assume that the distances are metric and that the function f is positive definite. An example of such a function is $f(d) = \exp(-\theta d)$, $\theta > 0$. According to (1), the conditional probability that B_j is a cure given that B_i is a cure declines from 1 to p as d_{ij} increases from 0 to ∞. It follows from (1) that:

$$prob(B_i \cap B_j) = p^2 + p(1-p)f(d_{ij})$$

In general, it is not possible to find p(S) from the univariate and bivariate marginal probabilities. It is, however, possible to place a lower bound on p(S). For arbitrary events A_i, i = 1, 2, ..., n, Gallot (1966) showed that:

$$\text{prob}\left(\bigcup_{i=1}^{n} A_i\right) \geq \sup_{c} \; (c' \; P_1 \; P_1' \; c) \; / \; (c' \; P_2 \; c) \qquad (2)$$

where $P_1 = (\text{prob}(A_1) \; \ldots \; \text{prob}(A_n))'$ and $P_2 = [\text{prob}(A_i \cap A_j)]$, i,j = 1, 2, ..., n. In terms of our model, this implies that:

$$p(S) \geq \sup_{c} \; [1 + p \; (1-p) \; (c' \; F \; c)/(c' \; P \; c)]^{-1} \qquad (3)$$

where $F = [f(d_{ij})]$, i,j = 1, 2, ..., n, and P is an n-by-n matrix with all elements equal to p^2. Provided that F is non-singular, it can be shown that:

$$\sup_{c} \; (c' \; P \; c) \; / \; (c' \; F \; c) = p^2 \; e' \; F^{-1} \; e \qquad (4)$$

where e is an n-vector of 1's. A similar result involving a generalized inverse of F follows from Kounias (1968) in the case where F is singular. For an arbitrary set A, we will denote $e' \; F^{-1}$ e by V(A). Since the lower bound on p(A) is an increasing function of V(A), different sets of species can be compared in terms of V(A), with larger values corresponding to greater lower bounds on

option value.

If $f(d_{ij}) = 0$ for all i, j (i.e., all pairs of species are independent), then V(A) is equal to the number of species in A. If $f(d_{ij})$ approaches 1 for all i, j (i.e., all pairs of species are perfectly dependent), then V(A) approaches 1. As discussed in the following section, while it is possible that V(A) can exceed n, we conjecture that for a reasonably constrained family of functions f V(A) lies between 1 and n. In this sense, V(A) can be interpreted as the effective number of species in A.

3. CONNECTIONS TO DIVERSITY

The most comprehensive treatment of the concept of diversity (as used in this paper) is by Weitzman (1992), who laid out three natural requirements of a diversity measure. Let S and T be two sets of species and let D(A) be a measure of the diversity of a set A. The first requirement of D(A) is that, if $S \subset T$, then $D(S) \leq D(T)$. This requirement, which is called monotonicity in species, says that diversity cannot be reduced by the addition of a species. The second requirement is that, if s_o is a single species, then $D(S \cup s_o) = D(S)$ if and only if $d_{oi} = 0$ for some $s_i \in S$. This requirement, which Weitzman (1992) refers to as twinning, says that diversity is not increased by the addition of a species if it is identical (in the sense of being at 0 distance) to a species already in the set. The third requirement is that, for a one-to-one mapping of S onto S′ such that $d_{ij} \leq d_{i',j'}$, with at least one

strict inequality, $D(S') > D(S)$. This requirement, which is called monotonicity in distance, says that diversity is increased by an unambiguous increase in the distances between species.

These three requirements do not define a unique measure. Weitzman (1992) proposed a diversity measure by defining the distance between a single species s_o and a set of species S as:

$$d(s_o, S) = \min_{s_i \epsilon S} (d_{oi})$$

The diversity measure is then given by the recursion:

$$W(S) = \max_{s_i \epsilon S} (W(S - s_i) + d(s_i, S - s_i)) \qquad (5)$$

where $S - s_i$ is the set formed by omitting s_i from S and where $W(S) = 0$ if S contains a single species. This measure can be motivated in the following way. Suppose that the diversity of the set $S \cup s_o$ is equal to the diversity of S plus $d(s_o, S)$. This allows the diversity of a set to be calculated by building it up one species at a time. Unfortunately, the result of this calculation depends on the order in which the species are added in building up S. This ambiguity is removed by imposing the maximization in (5). While W(S) satisfies monotonicity in species and twinning, it does not satisfy strict monotonicity in distance. That is, while W(S) cannot be reduced by increasing pairwise distances, it is possible to increase certain distances and leave W(S) unchanged.

Let $S = (s_1, s_2, \ldots, s_n)$ with:

$$V(S) = \sup_{c_n} (c_n' P_n c_n) / (c_n' F_S c_n)$$

in obvious notation. Let $S' = S \cup s_{n+1}$ and let $c^*_{n+1} = (c_n' \; 0)'$. Since:

$$\begin{aligned} V(S) &= \sup_{c_n} (c^*_{n+1}' P_{n+1} c^*_{n+1}) / (c^*_{n+1}' F_{S'} c^*_{n+1}) \\ &\leq \sup_{c_{n+1}} (c_{n+1}' P_{n+1} c_{n+1}) / (c_{n+1}' F_{S'} c_{n+1}) \\ &= V(S') \end{aligned}$$

$V(S')$ cannot be smaller than $V(S)$. This ensures that $V(A)$ is monotone in species.

Let S and S' be as above with $d_{1,n+1} = 0$. This means that the first and last rows and columns in $F_{S'}$ are identical. For any $(n+1)$-vector c_{n+1}:

$$\begin{aligned} &(c_{n+1}' P_{n+1} c_{n+1}) / (c_{n+1}' F_{S'} c_{n+1}) \\ &\quad = (c_n' P_n c_n) / (c_n' F_S c_n) \end{aligned} \tag{6}$$

where the first element of c_n is equal to the sum of the first and last elements of c_{n+1} and all other elements of c_n are equal to the corresponding elements of c_{n+1}. It follows that $V(S')$, which is the supremum over c_{n+1} of the lefthand-side of (6), is equal to $V(S)$, which is the supremum over c_n of the righthand-side of (6). This

ensures that V(A) satisfies twinning.

In general, V(A) is not monotone in distance. That is, it is possible to construct a positive definite matrix F with positive elements such that $e' F^{-1} e$ does not increase monotonically with a decrease in one of the off-diagonal elements of F (e.g., Eaton, 1992). It is possible to show in the 3-by-3 case that this result - which is connected to the comparison of experiments - cannot hold if all partial correlations are positive. Since the triangle inequality ensures that this condition is met for $f(d) = \exp(-\theta d)$, we conjecture that V(A) is monotone in distance and lies between 1 and n for this choice of f.

4. EXAMPLE

Rodman (1991) discussed the taxonomy of 26 species of plants that produce glucosinolates (sulfur-containing compounds related to mustard oils that have been identified as potential cancer-fighting agents). The species are listed in Table 1. Rodman (1991) used principal coordinate analysis (Gower, 1966) to display these species in 2 dimensions. Figure 1 shows the locations of the species with respect to the first two principal coordinates. For convenience, we will use the distances between these projections as the pairwise distances between the species.

Suppose that $f(d) = \exp(-\theta d)$ and suppose that we wish to compare the entire set of 26 species (S_1) to the set consisting of FLA, RES, and GER (S_2). If $\theta = 0.1$, then $V(S_1) = 2.62$ and $V(S_2) =$

2.32. Loosely speaking, the slow decline of f with d ensures that the dissimilarity effect (which favors S_2) is stronger than the size effect (which favors S_1) and the two sets are nearly equivalent in terms of the lower bound for option value. Because $S_2 \subset S_1$, monotonicity in species ensures that $V(S_1) \geq V(S_2)$.

If $\theta = 0.5$, then $V(S_1) = 10.55$ and $V(S_2) = 3.00$. In this case, the decline of f with d is more rapid, the size effect is more important, and S_1 is strongly preferred to S_2.

5. DISCUSSION

It is worth restating the practical problem addressed in this paper. To allocate scarce resources to the conservation of biological diversity in an effective way, it is necessary to measure biological diversity. One approach to constructing a measure of biological diversity is through guesswork - that is, by choosing a measure that seems to make sense (e.g., Eiswerth and Haney, 1992). A better approach is that followed by Weitzman (1992): namely, to set out some properties that a diversity measure should have and to identify measures that satisfy these properties. The main weakness of this approach is that the required properties are to some extent arbitrary, reflecting some general notion of what diversity means.

In this paper, we have taken a different approach. Instead of assuming a preference for diversity <u>per se</u>, we have developed a simple model relating the distances between species to a

nonparametric lower bound for option value. This model does little more than capture the assumption that the tendency to share characteristics is stronger for closely related species. Under this model, different sets of species can be compared (under a minimax criterion) using the simple measure V(A), which is a function of the distances between species. Provided that our conjecture regarding monotonicity in distance is true, it is at least mildly surprising that V(A) satisfies the three main requirements for a diversity measure. For example, if we had begun with the weaker lower bound for p(S) given by Chung and Erdös (1952) (which corresponds to Gallot's Inequality with c = e), the resulting measure would not have satisfied these requirements.

It should be emphasized that the issue addressed in this paper - namely, the measurement of biological diversity - is only one part of the problem of managing species conservation. One model of conservation decision-making (into which a measure such as V(A) can be embedded) was outlined in Solow, et al. (1992). Polasky and Solow (1992) outlined a more comprehensive model in which diversity is only one element of the value of species.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contributions of J. Broadus, M. Eaton, F. Fisher, W. Smith, R. Solow, Y. Tong, N. Tongring, and M. Weitzman to this paper.

REFERENCES

Brown, G.M. and Goldstein, J.H. (1984). A model for valuing endangered species. Journal of Environmental Economics and Management, 11, 303-309.

Chung, K.L. and Erdös, P. (1952). On the application of the Borel-Cantelli lemma. Transactions of the American Mathematical Society, 72, 179-186.

Eaton, M.L. (1992). A group action on covariances with applications to the comparison of linear normal experiments. Unpublished mimeo. Department of Statistics, University of Minnesota.

Eiswerth, M.E. and Haney, J.C. (1992). Allocating conservation expenditures across habitats: Accounting for inter-species genetic distinctiveness. Ecological Economics, in press.

Gallot, S. (1966). A bound for the maximum of a number of random variables. Journal of Applied Probability, 3, 556-558.

Gower, J.C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 29, 115-129.

Kounias, E.G. (1968). Bounds for the probability of a union, with applications. Annals of Mathematical Statistics, 39, 2154-2158.

Pielou, E.C. (1975). Ecological Diversity. Wiley Interscience, New York.

Polasky, S., Solow, A.R., and Broadus, J.M. (1992). Searching for uncertain benefits and the conservation of biological diversity. Journal of Environmental Economics, submitted.

Rodman, J.E. (1991). A taxonomic analysis of glucosinolate-producing plants, Part 1: Phenetics. Systematic Botany, 16, 598-618.

Solow, A.R., Polasky, S., and Broadus, J.M. (1992). On the measurement of biological diversity. Journal of Environmental Economics and Management, in press.

Vane-Wright, R.I., Humphries, C.J., and Williams, P.H. (1991). What to protect? - Systematics and the agony of choice. Biological Conservation, 55, 235-254.

Weitzman, M.L. (1992). On diversity. Quarterly Journal of Economics, 363-405.

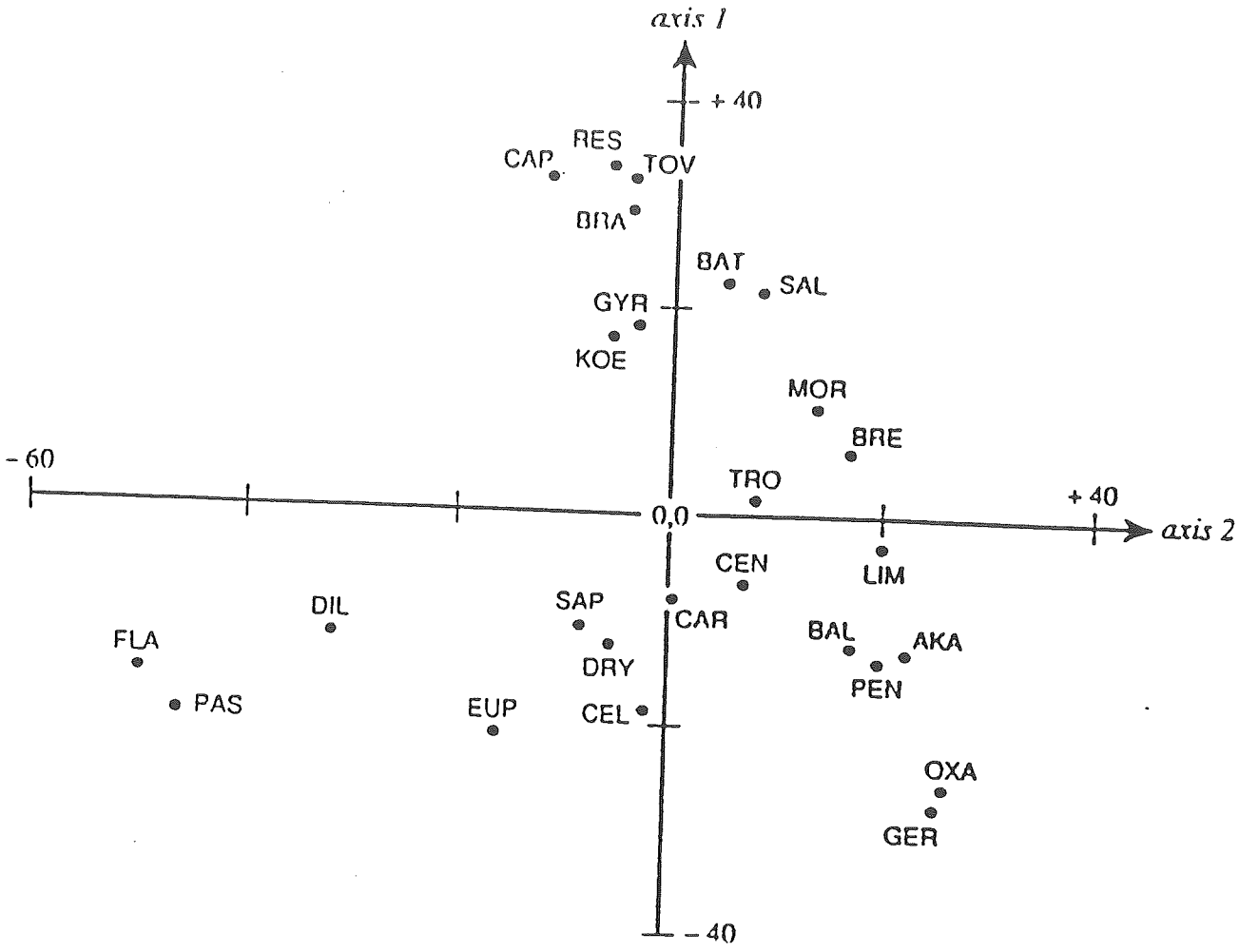
TABLE 1

Taxa of glucosinolate-producing plants and putative relatives.

Taxon	Code	Taxon	Code
Akaniaceae	AKA	Tovariaceae	TOV
Bataceae	BAT	Tropaeolaceae	TRO
Brassicaceae	BRA	Balsaminacea	BAL
Bretschneideraceae	BRE	Celastraceae	CEL
Capparaceae	CAP	Centrospermae	CEN
Cariacaceae	CAR	Dilleniaceae	DIL
Drypetes	DRY	Euphorbiaceae	EUP
Gyrostemonaceae	GYR	Flacouticeae	FLA
Limnanthaceae	LIM	Geraniaceae	GER
Moringaceae	MOR	Koeberliniaceae	KOE
Pentadiplandraceae	PEN	Oxalidaceae	OXA
Resedaceae	RES	Passifloraceae	PAS
Salvadoraceae	SAL	Sapindaceae	SAP

FIGURE CAPTION

Figure 1. Locations of the species in Table 1 in terms of the first two principal axes (from Rodman, 1991).


axis 1
+ 40
RES
CAP
TOV
BRA
BAT
SAL
GYR
KOE
MOR
BRE
- 60
TRO
+ 40
0,0
axis 2
CEN
LIM
DIL
SAP
CAR
FLA
BAL
AKA
DRY
PEN
PAS
EUP
CEL
OXA
GER
- 40