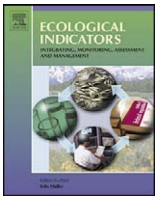
Ecological Indicators 11 (2011) 745–748

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



Short communication

Attempted tests of the surrogacy value of the ED environmental diversity measures highlight the need for corroboration assessment of surrogacy hypotheses

Daniel P. Faith

The Australian Museum, 6 College St., Sydney 2010, NSW, Australia

A R T I C L E I N F O

Article history:
Received 18 July 2009
Received in revised form 22 June 2010
Accepted 29 June 2010

Keywords:
Environmental diversity
Biodiversity surrogates
ED
Corroboration
Philosophy of science
Ordination

A B S T R A C T

A recent paper in this journal (Hortal et al., 2009) claimed to have evaluated the ED biodiversity surrogates methods of Faith and Walker (1994, 1996), and to have provided evidence for poor performance of the continuous ED method. In fact, their study neither used nor evaluated the continuous ED method. Here, I document their misrepresentation. I then discuss some constructive lessons emerging from their study and other recent studies that have attempted tests of ED surrogacy value. The need to consider the actual degree of support that observed evidence provides for a hypothesis about surrogacy raises general issues for evaluations of indicators' performance, and suggests a greater role for corroboration assessments. Guidelines for achieving this cover three aspects of surrogates testing: experimental design of tests, ongoing corroboration assessment of evidence produced by tests, and accumulation of lessons learned from multiple test studies over time.

1. Introduction

Biodiversity surrogates strategies use available species and environmental data to make inferences about general biodiversity patterns. Effective surrogates not only will act as a proxy for other known variation, but also will use the "known" to speak for the "unknown" – indicating patterns of geographic distribution for species not yet known to science. Surrogates therefore will work best when they reflect underlying ecological and evolutionary processes common to many species. Faith and Walker (1994, 1996; see also Faith, 2003; Faith et al., 2004) proposed "ED" biodiversity surrogates methods, based on the idea that environmental gradients explain the distribution of many different species. The rationale is that the degree of "environmental diversity" represented by a set of localities will indicate its relative species diversity. ED links species and environmental variation through an ecological pattern/process model that assumes general unimodal "responses" of species to environmental gradients (for background, see Faith et al., 1987). ED calculations are applied to localities positioned in an environmental space or ordination. While any environmental space in principle might be used, species and environmental data for localities typically are used in combination to derive the space, using multivariate methods compatible with a model of unimodal response. As a first step, compositional dissimilarities that are robust under the unimodal response model (see Faith et al., 1987) are calculated among sampled localities. Then, environmental variables may be used in a regression that predicts dissimilarity values (GDM; Ferrier et al., 2007, 2009), allowing ED calculations to use dissimilarities for all pairs of localities (Faith and Ferrier, 2002; Faith et al., 2004). These dissimilarities may be used directly in ED calculations. Alternatively, multidimensional scaling of the dissimilarities can position the localities in an environmental space for ED calculations (Faith and Walker, 1996; Faith and Ferrier, 2002; Faith, 2003; Faith et al., 2004).

The assumption of a general unimodal response model directly leads to the use of p-median (and related) optimization criteria as the basis for calculating ED values of selected sets of localities. A p-median criterion seeks to minimise the sum of distances from each "demand point" to its nearest selected site. For example, a p-median selection of a set of new hospital locations will minimise the average travel distance from each house (demand point) to its nearest hospital.

ED critically depends on the definition of sites and demand points in environmental space. The "continuous" form of ED refers to the case where the demand points are hypothetical points distributed uniformly throughout the continuous environmental space, and the "discrete" form of ED simply defines each of the candidate sites as a demand point (Faith and Walker, 1994, 1996). The continuous version of ED has been recommended (Faith and Walker, 1996; Faith, 2003; Faith et al., 2004) because discrete ED

E-mail address: dan.faith@austmus.gov.au.

1470-160X/$ – see front matter © 2010 Published by Elsevier Ltd.
doi:10.1016/j.ecolind.2010.06.019

may be sensitive to clumped or redundant distribution of localities in environmental space. Further, Faith and Walker (1996) proved, for continuous ED, that selection of sites using *p*-median will maximise the number of represented species, under the unimodal response model.

The link to unimodal response models also has guided the development of practical extensions of ED, including integration of variable species richness across the space, and use of probabilities rather than all-or-nothing selection of sites (Faith and Walker, 1996; Faith et al., 2004). The ED methods have had steady application (particularly for indicating expected gains in species representation from potential new survey sites, Funk et al., 2005; Ferrier et al., 2007; see also Juutinen et al., 2008; Hortal and Lobo, 2005). However, the ED methods also have generated controversy (e.g., Araújo et al., 2001, 2003, 2004; Faith, 2003; Faith et al., 2004). Faith (2003; see also Faith et al., 2004) suggested that studies reporting poor ED performance might be explained by specific properties of those particular studies, including sampling biases, poorly estimated environmental space, and the use of discrete rather than continuous ED.

The recent paper in this journal by Hortal et al. (2009) appears therefore to be highly relevant to these debates, in claiming evidence for relatively *poor* performance of continuous ED. However, Hortal et al. did not, contrary to their claims, evaluate continuous ED. While Hortal et al. did create demand points uniformly distributed in their environmental space, they then selected *demand points* instead of proceeding to select *sites* based on ED's *p*-median criterion (analogous to selecting the houses rather than the hospital locations). Their application of the *p*-median in effect tried to minimize the total distance from all demand points to their nearest selected *demand point* (not site). A *site* then was interpreted as selected only if it was the site closest to a selected demand point. This novel procedure is not equivalent to *p*-median's selection of *sites* to minimise sum of distances of demand points to nearest selected sites. Their calculations consequently fail to capture those key properties of continuous ED that promote its ability to maximise species representation. Indeed, their procedure totally excludes from selection sites that might be very attractive sites under the true continuous ED method. For example, in their Fig. 1 example, site 2 (Fig. 1d) never can be selected using their method, because it is not closest to any possible selected demand point. Yet, true continuous ED could select that site, and so represent species associated with the four surrounding demand points. The fact that the site is not even a candidate for selection by their procedure highlights the non-equivalence to continuous ED.

The non-equivalence to continuous ED is revealed by other properties of their procedure. The fact that sites are only selected indirectly, through selection of demand points, means that two selected demand points may be well-separated in space, yet their corresponding nearest-sites could be closer together. Picture two demand points on either end of a line segment in environmental space, with each of their closest sites positioned near the middle of the segment. Their method therefore could indirectly select two sites that were nearly identical. This reduces the capacity of a selected set of sites to effectively represent the entire space. This property of their procedure contrasts with the key property of the actual continuous ED method, which focuses directly on selecting those sites that complement other selected sites.

These properties of their procedure make it clear that Hortal et al. (2009) did not, as claimed, apply and evaluate the continuous ED method of Faith and Walker (1994, 1996). This casts doubt on their conclusion that "our results support findings that ED has only limited value as a surrogate for biodiversity."

2. When does the evidence strongly support a surrogacy hypothesis?

The Hortal et al. (2009) study did not provide its claimed contribution to the ED debates. However, their study does provide, as the latest in a series of ED evaluations, some strong impetus to consider how we should go about testing hypotheses about surrogacy. A critical issue is how and when some supposed positive evidence for a given surrogacy hypothesis can be interpreted as providing strong support for that hypothesis.

In the Hortal et al. study, as in previous studies (Araújo et al., 2001, 2004), evidence for a hypothesis that ED "has only limited value as a surrogate" was given by an observed poor recovery of species patterns. Such evidence is a discouraging result for any surrogate strategy, but we might be hesitant to draw firm conclusions if the properties of the study suggest that observed poor recovery could arise even for a very good surrogate strategy. I noted above that poor ED performance in some previous studies might be explained by sampling biases, by poorly estimated environmental space, and by the use of discrete rather than continuous ED (Faith, 2003). We can add to this list the possibility that an observed poor performance might result simply from substituting a different procedure for the actual ED method.

Faith (2003; see also Faith et al., 2004) suggested that such concerns can be handled constructively through an approach called corroboration assessment. Popperian corroboration is based on assessing whether apparent evidence in favour of a hypothesis is "improbable" without the hypothesis (Popper, 1959). Improbability simply means that the evidence cannot easily be explained, or accounted for, by other alternative explanations ("explained away"). Effective corroboration assessment therefore depends on an obligation to put forward any "background knowledge" (Popper, 1959) that would suggest that the observed evidence is "probable" (has a good chance of being observed) even without considering the hypothesis (for discussion, see Faith and Cranston, 1992; Faith, 2006, 2007).

Popperian corroboration for a hypothesis about ED surrogacy value therefore calls for some assessment that the supposed evidence for the hypothesis is otherwise "improbable" – it seems unlikely that we would have observed such good evidence through other causes. Any such assessment of improbability should be found *in spite of* genuine attempts to identify background knowledge that can provide alternative explanations of the observed evidence (exposing the evidence as "probable" based on other possible explanations).

The debates about the nature of evidence for surrogacy hypotheses have provided a context for discussion of corroboration assessment in the *Stanford Encyclopedia of Philosophy* (Faith, 2007):

> "For biodiversity surrogates, a common hypothesis is that the pattern of species "turnover" over different geographic areas for one taxonomic group will indicate the pattern for all biodiversity. Good evidence for the surrogacy hypothesis is typically claimed when the pattern for the surrogate taxonomic group is congruent with that of some target set of taxa. However, on many occasions such evidence can be explained away as probably arising simply because of a shared bias in the geographic sampling of the surrogate and target taxonomic groups (for review, see Faith, 2003). The evidence based on congruence can be explained away as a probable result even without the hypothesis. Based on such evidence, corroboration for the surrogacy hypothesis is low."

That example illustrates how what appears to be good evidence for surrogacy can suddenly appear weak in light of an alternative explanation for the test result.