Ecological Indicators 11 (2011) 896–901

Contents lists available at ScienceDirect

Ecological Indicators

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

Evaluation of five clustering algorithms for biodiversity surrogates

Rafi Kent*, Yohay Carmel

Department of Environmental Engineering, Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 14 September 2010
Received in revised form
29 November 2010
Accepted 5 December 2010

Key words:
Clustering algorithm
Conservation planning
Mt. Carmel

ABSTRACT

Conservation planning requires knowledge of the distribution of all species in the area of interest. Surrogates for biodiversity are considered as a possible solution. The two major types are biological and environmental surrogates. Here, we evaluate four different methods of hierarchical clustering, as well as one non-hierarchical method, in the context of producing surrogates for biodiversity. Each clustering method was used to produce maps of both surrogate types. We evaluated the representativeness of each clustering method by finding the average number of species represented in a set of sites, one site of each domain, which was carried out with Monte-Carlo permutations procedure. We propose an additional measure of surrogate performance, which is the degree of evenness of the different domains, e.g., by calculating Simpson's diversity index. Surrogates with low evenness leave little flexibility in site selection since often some of the domains may be represented by a single or very few sites, and thus surrogate maps with a high Simpson's index value may be more relevant for actual decision making. We found that there is a trade-off between species representativeness and evenness. Centroid clustering represented the most species, but had very low values of evenness. Ward's method of minimum variance represented more species than a random choice, and had high evenness values. Using the typical evaluation measures, the Centroid clustering method was most efficient for surrogate production. However, when Simpson's index is also considered, Ward's method of minimum variance is more appropriate for managers.

1. Introduction

Conservation of biodiversity requires extensive knowledge of the distribution of a myriad of species. Such knowledge is scarce, and collecting all necessary data is often prohibitively costly. One widespread solution is to use surrogates for biodiversity (e.g. Belbin, 1993; Faith and Walker, 1996; Ferrier and Watson, 1997). There are two types of biodiversity surrogates. In biological surrogates, the distribution of a taxonomic group is used to predict distribution patterns of other groups (Dobson et al., 1997; Reyers et al., 2001; Kati et al., 2004). Environmental surrogates are classifications of an area into land parcels with similar physical characteristics (Leathwick et al., 2003; Bonn and Gaston, 2005; Rodrigues and Brooks, 2007). An ongoing debate concerns the effectiveness of surrogates in predicting species assemblages (Faith and Walker, 1996; Rodrigues and Brooks, 2007). Ferrier and Watson (1997) proposed two ways to quantitatively assess the effectiveness of surrogates: (1) based on the number of species represented by a set of sites selected for conservation using different surrogates, and (2) on the level of correlation between the spatial structure of a surrogate and of the taxonomic group of interest.

A common approach for producing surrogates for biodiversity is classification with cluster analysis (Faith and Walker, 1996; Trakhtenbrot and Kadmon, 2005). Clustering can be conducted with various partitioning or agglomerating methods (Everitt, 1993; Legendre and Legendre, 1998), based on similarity or dissimilarity measures. Several arbitrary decisions are made during this process, regarding the similarity measures, number of classes and the specific clustering algorithm to be used (Everitt, 1993). These decisions may largely affect the resulting surrogate map.

Here, we evaluate five clustering methods commonly used in conservation planning, representing three different approaches to clustering, average based methods (Average, Centroid and Ward's minimum variance), object based methods (i.e. furthest neighbor) and non-hierarchical classification (k-means). This study attempts to take steps towards understanding surrogacy in two directions. First, we quantitatively evaluate the efficiency of different clustering methods for surrogate production. We assess both biological and environmental surrogates, by applying different algorithms to the same dataset. The second element of this study adds a new aspect to the evaluation of surrogate performance, measuring the evenness of surrogate classes, in addition to their species representativeness. Surrogates are tools for planning reserve networks; classes covering very small areas are more difficult to incorporate into such networks. Surrogates with a low evenness values are characterized by few dominant classes, and other classes occupying a negligible area (Fig. 1b). In surrogates with a high evenness value,

* Corresponding author. Tel.: +972 4 8293471; fax: +972 4 8293471.
E-mail addresses: rkent@tx.technion.ac.il, rkent@technion.ac.il (R. Kent).

1470-160X/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ecolind.2010.12.005

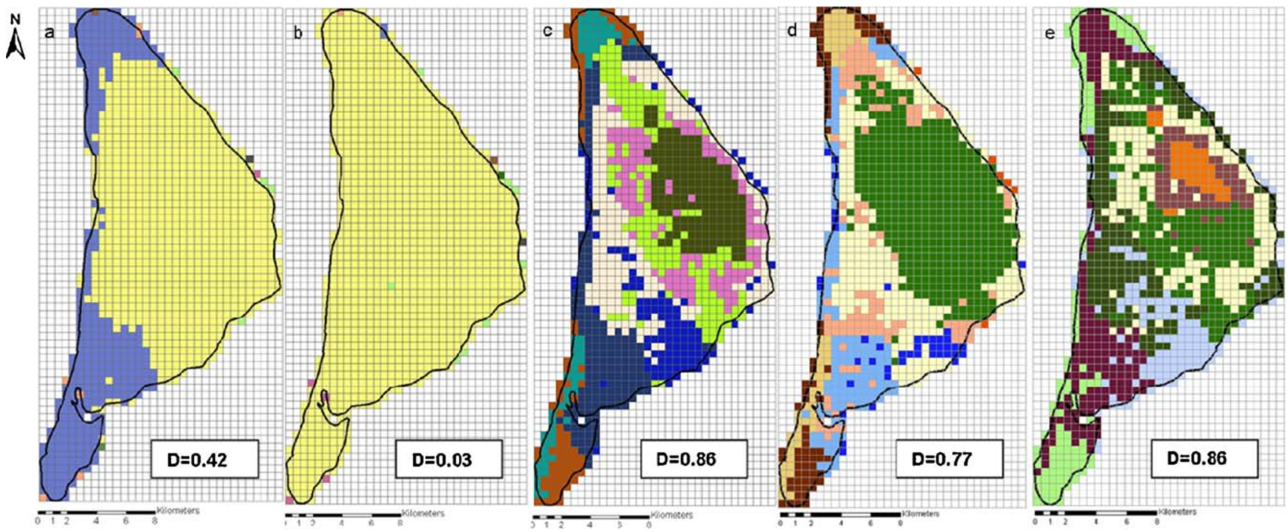

Fig. 1. Biological surrogates with eight classes, based on woody species distribution, produced with five different clustering algorithms: Average, Centroid, Ward's method, complete linkage and *k*-means clustering from left to right, respectively. *D* is Simpson's diversity index.

each class occupies a substantial part of the area. Such surrogate maps allow managers more flexibility in choosing areas for conservation and thus, more considerations may be taken into account, such as development, connectivity and land costs (Fig. 1). We use Simpson's diversity index as a measure of evenness.

2. Methods

2.1. Study area

Mt. Carmel in northern Israel has an area of ca. $\sim 280\,\text{km}^2$ and mean elevation of 220 m (Fig. 2). The climate is eastern Mediterranean, with mean annual rainfall of $650\,\text{mm year}^{-1}$ and temperature averages ranging between 11 °C in January and 24 °C in August. The most common soil types are Tera Rossa and Rendzina. Vegetation is comprised eastern Mediterranean scrubland consisting of structurally rich and diverse vegetation communities (Naveh and Dan, 1973; Le Honerou, 1981; Naveh and Kutiel, 1986). These landscapes, commonly referred to as vegetation mosaics, are highly heterogeneous at a broad range of spatial scales, ranging between grain size as small as a few meters to landscape level scales (Naveh, 1975; Shoshany, 2000; Bar Massada et al., 2008). The fine-grained mosaic is characterized by woody patches, herbaceous clearings, exposed rock and bare ground (Perevolotsky et al., 2002). A grid of 500 m × 500 m was superimposed on the area, dividing it into, 1145 cells. All analyses were carried out at this spatial scale.

2.2. Species distribution data

Presence–absence data for geophytes and woody plants were collected in 100 sampling sites distributed randomly in the entire study area, from October 2002 to May 2003 (see Carmel and Stoller-Cavari, 2006 for a complete description of fieldwork). Field data on each species, along with environmental parameters (Table 1), were used to create a habitat suitability map for each species. Field data were collected at 100 random sampling sites distributed across the study area, excluding urban and agricultural areas. Site size was 0.1 ha. Presence and absence of each woody plant and geophyte species were recorded at three sampling points within each site, with a distance of 20 m between sampling points. The sample area was $0.75\,\text{m}^2$. Habitat suitability maps were produced with logistic regression models (Guisan and Zimmermann, 2000) which were applied to the study area in the geographic

Fig. 2. A map of the study area, in the north of Israel.

Table 1
Environmental parameters used in the study for logistic regression models and environmental surrogates for biodiversity.

Variable	Source and spatial resolution
Normalized difference vegetation index—a measure of primary productivity	Produced from Landsat images, 25 × 25 m
Distance to nearest road	Produced from a vectoric roads map of the region, 500 × 500 m
Presence/absence of Terra Rossa soil in grid cell	Derived from polygons of soil association produced by the Agriculture ministry of Israel
Presence/absence of Rendzina soil in grid cell	
Presence/absence of woody vegetation cover in grid cell	Supervised classification of an aerial photograph of the region at 25 × 25 m
Average annual precipitation	Results of climatic models, produced at 25 × 25 m resolution
Mean daily temperature in the coldest month (January)	Results of climatic models, produced at 25 × 25 m resolution
Mean daily temperature in the hottest month (August)	Results of climatic models, produced at 25 × 25 m resolution
Mean geographic aspect of the slope in the grid cell	Derived from a 33 × 33 m resolution digital elevation model
Mean slope (in degrees) in the grid cell	Derived from a 33 × 33 m resolution digital elevation model
Mean elevation in the grid cell	Derived from a 33 × 33 m resolution digital elevation model

information system (GIS). Habitat suitability maps consisting of probabilities of occurrence, were transformed into binary (0/1) maps, with a threshold of 0.5. Statistically significant models were produced for 23 geophyte species and 37 woody species. These 60 distribution maps were used in further analyses (see Fig. 3 for example).

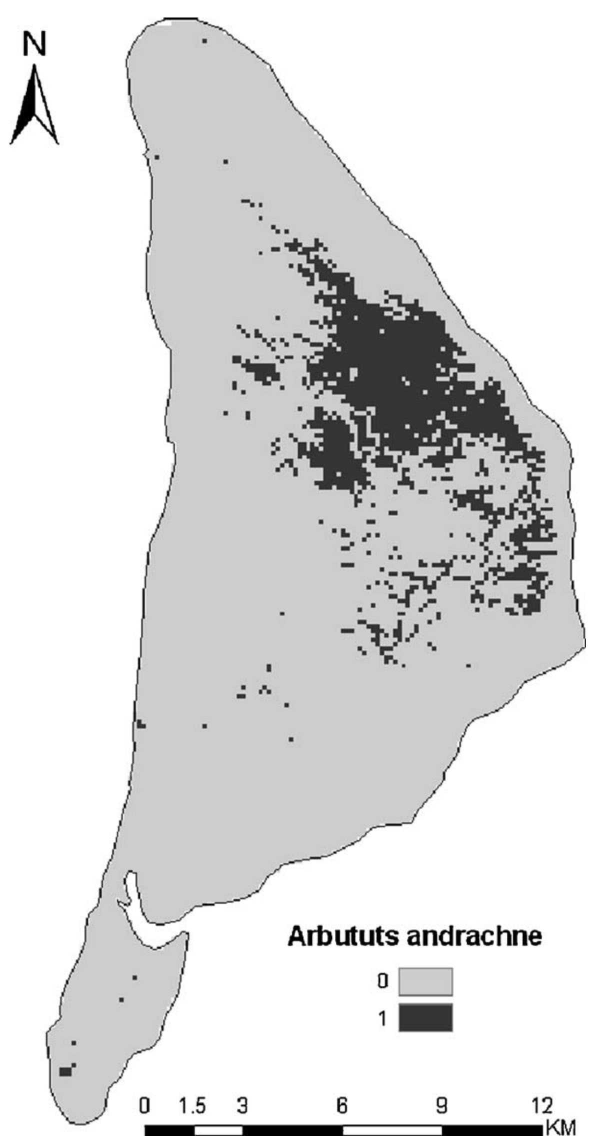

Fig. 3. Habitat suitability map for the tree species *Arbutus andrachne* on Mt. Carmel, northern Israel, produced with a logistic regression species distribution model.

2.3. Environmental data

Environmental parameters used in this study included measures of climate, soils, vegetation, topography and an anthropogenic disturbance index (Table 1). All topographic parameters were calculated from a digital elevation model of the study area. Parameters were chosen to represent independent influences, i.e., climate has a different effect than topography; primary productivity affects species composition differently than distance to road or soil type, etc. A total of 11 parameters were used, both for the logistic regression models (that produced species distribution maps), and for the environmental surrogates. All parameters were extracted from GIS layers and averaged over 500 m cells in order to fit the grid.

2.4. Surrogate production

Both types of surrogates, biological and environmental, were produced using five clustering algorithms: Average, Centroid and Ward's (average based approach), furthest neighbor (object based approach) and *k*-means (non-hierarchical classification) (Legendre and Legendre, 1998). We applied each clustering method to three schemes—three, eight and twelve domains. We used the two functional groups (geophytes and woody plants) alternately, as surrogates for each other. In total, we examined 45 different surrogates.

2.5. Biological surrogates

We calculated two similarity matrices for woody and geophyte species using Jaccard's similarity coefficient. The similarity matrices were then used to classify the 1145 cells into domains. We constructed three surrogate types: environmental surrogates, biological surrogates based on woody species distributions and biological surrogates based on geophyte species distributions. Each of these surrogate types was produced with three different domain numbers: three, eight and twelve, yielding nine combinations of surrogate type and the number of domains. Each such combination was produced using each of five different clustering methods: Average, Centroid, Ward's, complete linkage and *k*-means. In total we produced 45 surrogate maps.

2.6. Environmental surrogates

Environmental surrogates were produced in the same manner as the biological surrogates, using GIS layers of environmental parameters instead of species distributions (Table 1). All environmental parameters received the same weight in the analysis. We chose Gower's similarity coefficient instead of Jaccard's in order to produce environmental surrogates, since it allows incorporation of both continuous and binary parameters, such as presence/absence of a soil type or woody vegetation (Gower, 1971; Legendre and Legendre, 1998). In addition, when applied to binary data, Gower is equal to Jaccard, thus the two indices are compatible and interchangeable (Dunn and Everitt, 1982).

2.7. Performance evaluation

Performance of the five different types of clustering was compared in two ways: (a) species representativeness—the number of species that are represented when a single site of each class is chosen randomly and (b) Simpson's diversity index, in order to evaluate the evenness of the different domains within each surrogate map.

For species representativeness we recorded the cumulative number of species present in a set of sites, when a single site from each domain is chosen at random. Presence of species was determined by the distribution maps generated by habitat suitability models, as described above. Ten thousand random sets of sites were selected using a Monte-Carlo permutation procedure, one site from each domain. As a baseline for evaluating species representativeness, we used the permutation procedure to select completely random sets of sites, matching the clustering schemes (3, 8 and 12 sites) and recorded the average number of species represented in such sets over the 10,000 permutations. Environmental surrogates were evaluated using both taxonomic groups, while biological surrogates were evaluated against each other.

Simpson's diversity index was calculated for each set of surrogates, to test the diversity of the different classes.

To evaluate the different clustering algorithms, we gave the different algorithms a rank between 1 and 5, for their performance compared to the other methods for each surrogate. We ranked environmental surrogates according to their performance for each taxonomic group separately, and used the average rank for the final ranks. We ranked biological surrogates according to their performance on the other group. We gave each surrogate type a separate rank for each efficiency measure, and added them for total relative efficiency. In addition, we evaluated nine environmental surrogate maps of the entire flora of Israel, produced by Trakhtenbrot and Kadmon (2006) with three different algorithms (Average, Centroid and Ward's), and with 3 levels of partitioning (3, 8 and 12 classes). Trakhtenbrot and Kadmon report that the relative efficiency (representativeness) of the different algorithms was constant regardless of the number of classes. Average clustering performed best, followed by Centroid, and Ward's method was the least effective. Here, we calculated Simpson's index for these nine surrogate maps, and ranked the different maps accordingly.

3. Results

3.1. Species representativeness

The Monte-Carlo permutations revealed that Centroid clustering method was the most efficient algorithm in representing species, representing an average of 94% of the target species for the eight class scheme. The Centroid method was superior to the other methods in every combination of environmental surrogates, and in four of six combinations of biological surrogates. The Average clustering method was ranked highest in two combinations, representing an average of ~86% of the species, Ward's method was ranked third in species representativeness, representing an average of ~82% of the species for a scheme of eight classes. Complete linkage and *k*-means were the least effective methods for representing species with 79% and 81%, respectively. Although these are similar levels of representativeness as Ward's method, when compared to a set of randomly selected sites, both Complete linkage and *k*-means were the only methods that represented less species than such a random set of sites in one of the 9 surrogate combinations (Fig. 4). Fig. 4 shows the results the Monte-Carlo permutations of species representativeness for surrogates with eight classes. Surrogates with three and 12 classes showed a similar trend. All types of surrogates except one based on complete linkage and one based on *k*-means clustering represented more species than a random choice of the same number of cells from the grid (Fig. 4). We ranked the different clustering methods according to their relative efficiency, in order to create a single combined efficiency measure, of both representativeness and evenness, summarized in Table 2.

Table 2
Summary of the ranks of all different combinations of the number of classes and clustering algorithm used in the production process. Each algorithm was ranked 1–5 according to its relative efficiency in each clustering scheme (3, 8 and 12 classes), for each of the two performance evaluation measures (representativeness, Simpson's index). Values in the table are the sum of the ranks of each algorithm. Ranks have no units.

Clustering algorithm	Representativeness	Simpson's index
Algorithm ranking		
Average	32.5	18
Centroid	38.5	10
Ward's	22	43
Complete linkage	23.5	25
k-Means	19	39

3.2. Domains evenness

Fig. 1 shows that the Centroid method effectively produced only one large class across the entire area (Fig. 1b), and the remaining seven classes occupy a very small area on the perimeter. Four classes occupied only a single cell, and one class occupied two cells, with a value of Simpson's index of 0.03 (Fig. 5), apparently representing locations of rare species. Similarly, the Average method produced two large classes and six very small classes (Fig. 1a), with four classes occupying only a single grid cell, resulting in Simpson's index of 0.42 (Fig. 5). Ward's method produced eight effective classes (Fig. 1c), with a value of Simpson's index of 0.86 (Fig. 5) for the map in Fig. 1c, and so did complete linkage and *k*-means ($D = 0.77$ and 0.86, respectively). Ward's method produced surrogates with the highest Simpson's index value in seven out of nine cases (Fig. 5). We ranked the different clustering algorithms according to their Simpson's index values. Ranks given according to Simpson's index are summarized in Table 2.

4. Discussion

Evaluating the efficiency of surrogates for biodiversity is a complex task that requires knowledge of the entire biological diversity (Rodrigues and Brooks, 2007). A common solution is to use a group of known species as the target group, and to evaluate how well the surrogates represent them. Ferrier and Watson (1997) introduced a measure of efficiency, the species accumulation index, which measures how many species out of the entire pool are represented by a set of domains, and also compares it to a random selection of sites. Another measure they proposed was the Mantel correlation coefficient, which is the correspondence between domains and species distributions (Manly, 2007). These measures do not fully account for an important aspect of surrogacy, namely, its actual use for site selection, and additional information is needed, such as the evenness of the different domains. If domains are relatively evenly sized, higher flexibility is allowed for site-selection between various options, whereas if some domains are represented by very few pixels (as in the case of the centroind algorithm output), little flexibility is left to managers in the site selection process. Evenness is an intrinsic trait, thus it is an independent measure of surrogate quality.

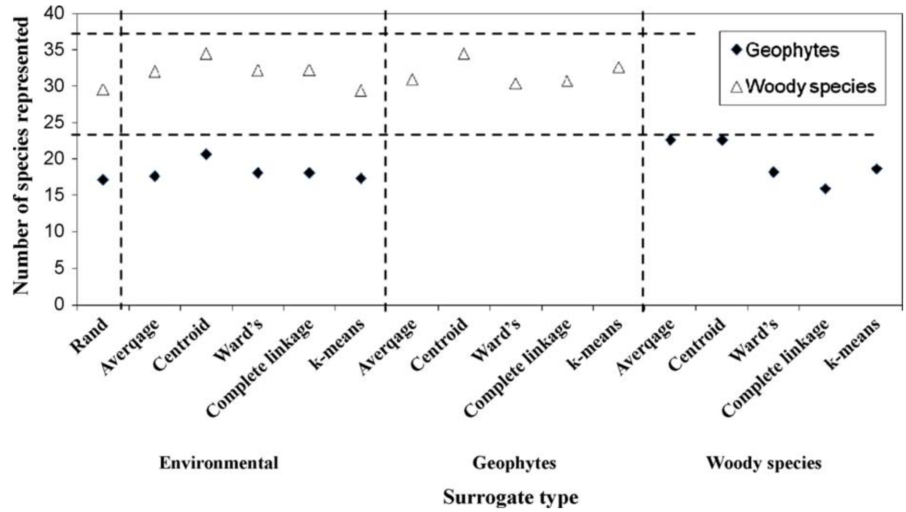

Fig. 4. Average number of species represented by a set of surrogates for the eight classes clustering scheme. Surrogates based on geophytes were tested with woody plants only, and vice versa. 'Rand' stands for a random selection of eight cells without classification, repeated 10,000 times. Values are average of 10,000 permutations of random selection of sites, one site per surrogate class.

The various algorithms could be divided into two groups: the Average and Centroid methods, which represented more species than the other algorithms, but had low values of Simpson's index, and Ward's, complete linkage and *k*-means, which represented less species but had relatively high Simpson's values. Due to the fact that we used modeled distributions to determine presence of species in selected sites, and since we had a relatively small number of species (23 and 37 species of geophytes and woody plants, respectively), there were multiple spatial configurations that represented most of the species. That was evident in the relatively small differences in species representativeness. However, applied to a larger species pool, even a 15% difference in the number of represented species may actually mean leaving dozens of species represented in a reserve network using a given method, or protecting them using another. Unfortunately, none of the algorithms we evaluated excelled in both species representativeness and domain evenness.

We used distribution data that we extrapolated from a dedicated survey as our biological dataset. Such extrapolation is bound to contain errors of omission and commission, i.e. to predict to absence of a species from a location where it exists, and to predict the presence of a species in a location from which it is actually missing. However, since we conducted the performance evaluation of the different methods on the modeled distributions, such omission and commission errors do not affect the results of our evaluation. For the purpose of planning an actual reserve network, it is necessary to validate such models, and assess their accuracy levels prior to proceeding with surrogate production.

Current conservation practice takes place at two levels, coarse and fine filters (Maddock and Du Plessis, 1999; Bonn and Gaston, 2005; Orme et al., 2005). Biodiversity surrogates are considered coarse filters, which are tools for capturing biodiversity in its broadest sense, including habitats, ecological processes and entire ecosystems, as well as individual species (Noss, 1990). Fine filters concern rare or endangered species, which could have been missed by coarse filter methods. Surrogates with domains that cover very small geographic areas (low Simpson Index value), such as those produced by the Average and Centroid clustering algorithms, may be difficult for use as coarse-filters, due to the local nature of some of the resulting domains.

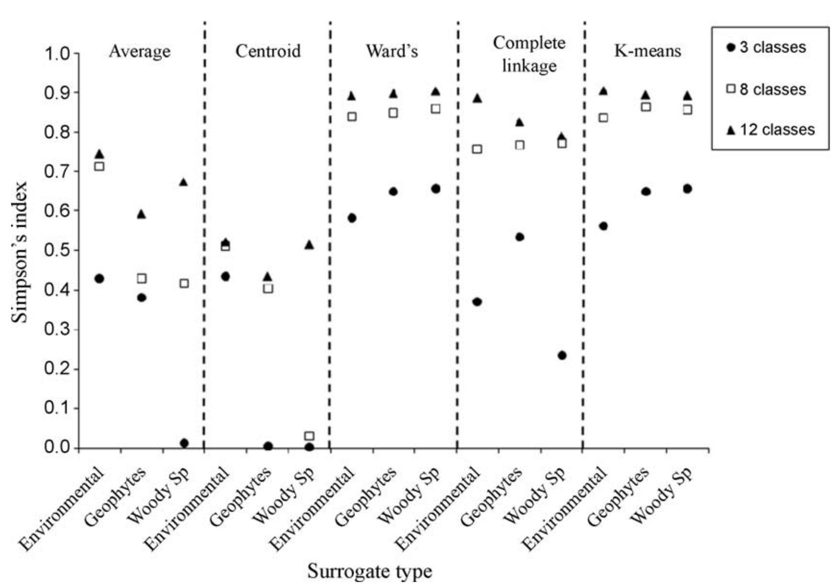

Fig. 5. Simpson's diversity index for each type of surrogate.

The effect of the clustering algorithm on the efficiency of surrogacy has been largely overlooked, and very few studies quantitatively evaluated alternative algorithms. Trakhtenbrot and Kadmon (2006) reported that the Centroid algorithm resulted in the best surrogates, which is similar to our results, when considering species representativeness alone. However, by calculating Simpson's diversity index in addition to species representativeness, we show that what seemed to be the better surrogate (Centroid-based cluster analysis), may be a less effective tool for biodiversity conservation from certain perspectives, although the number of species it represents may be larger. In general, our results point out that there is no 'optimal solution' to surrogacy, and each case should be considered separately. Trakhtenbrot and Kadmon (2006) reported that, when applying a weighting scheme similar to the weights in our study, the method that ranked highest in species representativeness of vascular plants in the entire state of Israel was Average, followed by Centroid, and that Ward's method was the least effective. Analyses of Trakhtenbrot and Kadmon's (2006) maps showed that Ward's method always produced maps with higher values of Simpson's index, and Average clustering was superior to Centroid in two of the three combinations. Our results, along with the analyses we conducted on Trakhtenbrot and Kadmon's (2006) data, imply that there is a trade-off between species representativeness and domain evenness.

Knight et al. (2008) reported the gap between conservation science and conservation actions. Our results indicate that a solution that may be considered scientifically superior may, in fact, be incompatible for managers. Fig. 1 is an example of the difference between scientifically 'superior' and management-compatible solutions. We conclude that scientific work on prioritization of areas for conservation should include a final step, which is mostly disregarded in the literature, i.e. looking at the resulting maps, and examining them through a field conservationist's eye.

Acknowledgements

This study was supported by Meron Ecological Center and the Open Land Institute of the Society for the Protection of Nature in Israel.

References

Bar Massada, A., Gabay, O., Perevolotsky, A., Carmel, Y., 2008. Quantifying the effect of grazing and shrub-clearing on small scale spatial pattern of vegetation. Landscape Ecol. 23, 327–339.

Belbin, L., 1993. Environmental representativeness—regional partitioning and reserve selection. Biol. Conserv. 66, 223–230.

Bonn, A., Gaston, K.J., 2005. Capturing biodiversity: selecting priority areas for conservation using different criteria. Biodivers. Conserv. 14, 1083–1100.

Carmel, Y., Stoller-Cavari, L., 2006. Comparing environmental and biological surrogates for biodiversity at a local scale. Israel J. Ecol. Evol. 52, 11–28.

Dobson, A.P., Rodrigues, J.P., Roberts, W.M., Wilcove, D.S., 1997. Geographic distribution of endangered species in the United States. Science 275, 553–1553.

Dunn, G., Everitt, B.S., 1982. An Introduction to Mathematical Taxonomy. Cambridge University Press, Cambridge.

Everitt, B.S., 1993. Cluster Analysis, third ed. John Wiley & Sons Inc., New-York.

Faith, D.P., Walker, P.A., 1996. Environmental diversity: the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. Biodivers. Conserv. 5, 399–415.

Ferrier, S., Watson, G., 1997. An Evaluation of the Effectiveness of Environmental Surrogates and Modelling Techniques in Predicting the Distribution of Biological Diversity. New South Wales National Parks and Wildlife Service.

Gower, J.C., 1971. A general coefficient of similarity and some of its properties. Biometrics 27, 857–874.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135, 147–186.

Kati, V., Devillers, P., Dufrene, M., Legakis, A., Vokou, D., Lebrun, P., 2004. Testing the value of six taxonomic groups as biodiversity indicators at a local scale. Conserv. Biol. 18, 667–675.

Knight, A.T., Cowling, R.M., Rouget, M., Balmford, A., Lombard, A.M., Campbell, B.M., 2008. Knowing but not doing: selecting priority conservation areas and the research-implementation gap. Conserv. Biol. 22, 610–617.

Le Honerou, H.N., 1981. Impact of man and his animals on Mediterranean vegetation. Ecosystems of the World, II. In: Di Castri, F., Goodall, D.W., Specht, R.L. (Eds.), Mediterranean Shrublands. Elsevier Scientific, Amsterdam.

Leathwick, J., Overton, J.M., Mcleod, M., 2003. An environmental domain classification of New Zealand and its use as a tool for biodiversity management. Conserv. Biol. 17, 1612–1623.

Legendre, P., Legendre, L., 1998. Numerical Ecology, second ed. Elsevier Science, Amsterdam.

Maddock, A., Du Plessis, M.A., 1999. Can species data only be appropriately used to conserve biodiversity? Biodivers. Conserv. 8, 603–615.

Manly, B.F.J., 2007. Randomization, Bootstrap and Mote Carlo Methods in Biology, third ed. Taylor and Francis Group, Boca Raton.

Naveh, Z., 1975. The evolutionary significance of fire in the Mediterranean region. Vegetation 29, 199–208.

Naveh, Z., Dan, J., 1973. The human degradation of Mediterranean landscapes in Israel. In: di Castri, F., Mooney, H.A. (Eds.), Mediterranean-Type Ecosystems, Origin and Structure. Springer-Verlag, New York.

Naveh, Z., Kutiel, P., 1986. Changes in the Mediterranean vegetation of Israel in response to human habitation and land use, The Earth in Transition. In: Woodwell, G.M. (Ed.), Patterns and Processes of Biotic Impoverishment. Cambridge University Press, Cambridge, pp. 259–296.

Noss, R.F., 1990. Indicators for monitoring biodiversity: a hierarchical approach. Conserv. Biol. 4, 355–364.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., Webster, A.J., Ding, T.S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennet, P.M., Blackburn, T.M., Gaston, K.J., Owens, I.P.F., 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436, 1016–1019.

Perevolotsky, A., Ettinger, E., Schwartz-Tzachor, R., Yonatan, R., 2002. Management of fuel breaks in the Israeli Mediterranean Ecosystem: the case of Ramat Hanadiv Park. J. Mediterr. Ecol. 3, 13–22.

Reyers, B., Fairbanks, D.H., Jaarsveld, A.S.V., Thompson, M., 2001. Priority areas for the conservation of South African vegettion: a coarse-filter approach. Divers. Distrib. 7, 79–95.

Rodrigues, A.S.L., Brooks, T.M., 2007. Shortcuts for biodiversity conservation planning: the effectiveness of surrogates. Annu. Rev. Ecol. Syst. 38, 713–737.

Shoshany, M., 2000. Satellite remote sensing of natural Mediterranean vegetation: a review within an ecological context. Progr. Phys. Geogr. 24, 153–178.

Trakhtenbrot, A., Kadmon, R., 2005. Environmental cluster analysis as a tool for selecting complementary networks of conservation sites. Ecol. Appl. 15, 335–345.

Trakhtenbrot, A., Kadmon, R., 2006. Effectiveness of environmental cluster anlysis in representing regional species diversity. Conserv. Biol. 20, 1087–1098.