

Determining species richness and endemism upon seamounts by utilising a hierarchical model system.

or

A hierarchical model for determining species richness at multiple spatial scales utilising endemism upon seamounts as a representative example.

Derek Tittensor, Ransom Myers, Karen Stocks

8th May 2004 v. 4.0.

1. Introduction and motivation

Seamounts are submarine topographical features that are linked with high biological diversity and an unusually large proportion of endemic species (Rogers, 1994), factors that make them important from scientific, conservation, and biodiversity standpoints. There remains an urgent need to assess species richness and species-area relationships in marine environments (Neigel, 2003) and the unique nature of seamounts make them prime candidates for further investigation. Unfortunately, given their enormous (>30,000; Rogers, 1994) number and variety, and the finite nature of the sampling resources available, some kind of targeting and allocation of the

aforementioned resources must be applied, with the intention of maximizing the return of scientific and policy-relevant knowledge upon the effort invested. We discuss a simple model structure that can be utilised in order to make preliminary estimates of the relative importance of some of the many parameters that determine a seamount's biological makeup, and may be of some aid when considering how best to sample these regions. The results of further sampling and data acquisition can then be fed back into this model in order to revise parameterisations and estimates, thus providing a coupled system in which empirical data can be used in order to improve the model, and model output can be used as a tool to aid with practical sampling issues. This model can then be used as the basis of a hierarchical structure that will allow estimation for the of species richness at many different scales.

2. Seamounts

Studies on seamounts appear to indicate that a high proportion of the species present upon them are endemic (e.g. Parin, 1991) in comparison to other oceanic regions of similar size. Even seamounts that are, geographically speaking, relatively closely spaced appear to have significant differences in community composition and large numbers of endemics (de Forges *et al.*, 2000). A question that invites further attention is whether these endemic species are really just that, or if they are 'false' artifactual endemics due to under-sampling of the surrounding benthos (de Forges *et al.*, 2000). This can only be addressed with further sampling.

A large number of trawls and dredges have provided extensive information about many different seamounts (Rogers, 1994), although the vast proportion have not been investigated and an enormous number of questions remain. For instance: what is the relationship between seamount isolation and endemism; how adequately do the samples reflect the underlying structure of the communities surveyed, and what are the biases of the different sampling methodologies (Rogers, 1994); what effect does the size of a seamount have on its species richness, and the number of endemic species that are present; what are the differences in community between structurally, geologically and oceanographically varying seamounts; are there significant differences between seamounts located in disparate regions within the same ocean, or between oceans; what effect does removing a proportion of the habitat from a seamount have on the species composition; how does this effect vary between seamounts of differing size and makeup; should sampling be divided among as many seamounts as possible in an attempt to quantify the differences between them, or should it be allocated to intensively census one or two locations in order to provide a more thorough understanding of a few selected sites; should sampling be repeated at evenly spaced temporal intervals in order to observe seasonal and longer time-scale fluctuations, in regions both undisturbed and those under the influence of anthropogenic disturbance.

The apparently high number of endemic species and relative isolation of seamounts suggest a comparison can be drawn with terrestrial island regions. MacArthur and Wilson's (1967) theory of island biogeography might thus provide clues as to the community structure, species interrelationships, effects of isolation, and immigration

patterns upon these regions, although care must be taken about applying such a broad theoretical hypothesis without first acknowledging the large- and small-scale differences in the physical and biological makeup of these ecosystems.

3. The single seamount model

This section describes a simple model for the number of endemic species upon a seamount. We suggest that a comparison can be drawn between seamounts and islands, in terms of the effects of isolation from source populations (the mainland/the continental shelf) and the proportion of endemic species to be found upon them. The theory of island biogeography (MacArthur & Wilson, 1967) hypothesises that the total number of species to be found upon an island is a power-function of the island's area. We might propose that the same holds for the total number of species upon a seamount, of which a proportion p are endemic:

$$E(e_s) = p(fA^z) \tag{1}$$

where e_s is the total number of endemic species upon seamount s , E is the expectation of this value, A is the total area of the seamount, and f and z are constants. We might also expect that the proportion of endemic species is a function of the isolation distance from the nearest source population (here defined to be another seamount, an island, or the continental shelf). We consider the isolation distance to saturate to an asymptote, after which the increased speciation effect of isolation is balanced by the distance being

greater than a characteristic maximum dispersal range x and therefore a reduced immigration rate. Hence we can replace p :

$$p = \frac{bD^y}{1 + bD^y} \quad 2$$

where D is the isolation distance and b and y are constants. Combining (1) and (2) we get:

$$E(e_s) = \frac{bD^y fA^z}{1 + bD^y} \quad 3$$

In this simple model, considering only isolation and size as important contributing factors, endemism can in a sense be considered as a species-area-isolation relationship (SAIR).

Determination of parameter values could utilize a number of sources: an obvious first step would be to assimilate the existing species data from seamounts and attempt to derive the values of best fit. An alternative route would be to use parametric distribution models that are appropriately constructed for the environment under study. We would suggest implementing the model within a Bayesian framework (Hillborn & Mangel, 2003) in order to utilise prior information. A number of candidate models and functional forms could be considered and an appropriate criterion (e.g. Bayesian Information Criterion) applied in order to determine the model of best fit (Johnson & Omland, 2004).

The broad scale and small number of parameters in the equation mean that it is obviously an over-simplification; among the factors that are not incorporated are topography, habitat heterogeneity, oceanographic conditions, geological makeup, depth, and many other features. However, we feel that the pressing need for further information points to the creation of a simple analytically tractable model in an attempt to shed some light upon these issues.

One point that we would like to note is that MacArthur & Wilson (1963) discuss a “radiation zone”, being the outermost dispersal limit of a particular taxon. This is what we call the characteristic maximum dispersal range, x , and is different for every individual species. It is certainly possible that seamounts that are further out than a certain isolation distance (depending upon the taxon present upon the seamount) contain a *smaller* number of endemic species than those further in, rather than the isolation function tending towards and asymptote. However, given the presumed greater dispersal distances within oceans, and the paucity of our knowledge on the long-term evolutionary dispersal processes, we felt it best to leave the isolation function in its current form. The measurement of the effects of isolation upon seamounts might provide some insight into whether they do indeed act as ‘stepping stones’ for trans-oceanic dispersal.

Equally, it is interesting to note that some seamounts that are not large or isolated appear to have a high proportion of endemic species (Koslow *et al.*, 2001). These deviations may be caused by differences in habitat type or other unknowns; there will almost certainly indeed be several differing sets of parameters for the SAIR that are

utilised for a number of different categories of seamount. Exceptional cases can be used to further study the causes of variability within seamounts and seamount chains that are not included within the model.

4. The hierarchical model

The ‘single seamount’ model can be fitted into a hierarchically organised model of species richness (as a function of endemism) at multiple spatial scales (see Figure 1). To begin with, the total number of species t_s that exist upon a particular seamount is described in equation 4:

$$t_s = e_s + \sum_{j=1}^{e_c} p_c(j) + \sum_{k=1}^{e_o} p_o(k) + \sum_{l=1}^{e_w} p_w(l) \quad 4$$

where e_c is the number of species endemic to seamount chain c within which seamount s is embedded (if the seamount is isolated and not part of a chain the $e_c = 0$), e_o is the number of species endemic to the ocean or sea within which the seamount is located, and e_w is the number of species that are present within more than one ocean and sea. p_c , p_o , and p_w are binary-valued parameters that define the presence or absence of a particular species (at the chain, ocean or world scale) upon seamount s , with the value 0 representing absence and 1 representing presence. The p values are dependent upon a number of factors, including (but not limited to) the size ratio between the seamount and the higher-level spatial structure within which it is embedded, the distance between seamounts (at the chain level), the distribution and patchiness of each species – once

more a link to the species-area relationship – and the diversity and partitioning of habitats upon seamount i . An appropriate practical method for determining the value of a p parameter for an individual species (barring a complete census of the seamount in question) might be to consider it as a probability distribution function that can be formulated from the parameters listed above and any others considered important.

The total number of species upon the n seamounts of the world is then:

$$\sum_{s=1}^n t_s = \sum_{s=1}^n e_s + \sum_{c=1}^r e_c + \sum_{l=1}^{e_o} q_o(l) + \sum_{m=1}^{e_w} q_w(m) \quad 5$$

where q_o is a binary function similar to p_o but representing the presence or absence of species endemic to ocean o upon *any* of the seamounts within a particular ocean and q_w is similarly valued for species that have a global marine presence. Note that the total number of species cannot be calculated simply by summing the total number of species from each individual seamount since many species may well be present upon more than one seamount.

The above equations can be thought of as two levels in a hierarchy; several layers could potentially be constructed between them (such as the number of species upon a chain of seamounts, or upon all the seamounts in a particular ocean). We now consider the next level ‘up’ in the hierarchy (where ‘up’ represents an increase in spatial scale, and ‘down’ the converse): the total number of species within the marine environment as a whole:

$$t_m = \sum_{s=1}^n e_s + \sum_{c=1}^k e_c + \sum_{o=1}^l e_o + e_w \quad 6$$

But how about the number of species within a particular ocean o ? This is obviously less than the total number of species in all oceans, but where does it fit in the hierarchy in relation to the total number of species upon all seamounts in all oceans? We can define hierarchical levels by saying that one level encompasses another if it contains the second level completely within itself, topographically speaking. Thus, since ocean o does not contain all n seamounts, and all n seamounts do not contain ocean o then these are at the same hierarchical level, and both contained by the total number of species within the marine environment as a whole. Equation (7) represents the total number of species within ocean o :

$$t_o = \sum_{s=1}^n c_s e_s + \sum_{c=1}^r c_c e_c + e_o + \sum_{m=1}^{e_w} r_w(m) \quad 7$$

where c_s and c_c are binary-valued functions that take the value 1 if, topographically speaking, $s \in o$ or $c \in o$ respectively, and zero otherwise. r_w is a binary-valued function that takes the value of 1 if a species that is present within more than one ocean/sea is contained within ocean o . Equation (6) can then be embedded within a global scale model of the total number of species within the biosphere:

$$t_w = t_h + t_l + t_t + t_a - t_{dup} \quad 8$$

where t_h is the total number of species within the hydrosphere, t_l is the total number of species within the lithosphere, t_t is the total number of terrestrial species, and t_a is the total number of atmospheric species, and t_{dup} is the total number of duplicated species that exist within more than one of the environments within the equation (such as amphibians).

5. Discussion

Equation (8) represents an alternative approach to May (1988) for determining the number of species present upon the planet on which we live; its hierarchical nature affords us an opportunity to combine the various data and estimates that exist into a cohesive whole, and represents a structural entity that is common within scientific endeavour (e.g. the Linnaean classification system) and is thus familiar. We have used endemism as the basis for determining species presence or absence; alternative functional forms could also be considered. Examining each hierarchical level within a probabilistic framework may give a predicted species richness value that can be statistically bounded with confidence limits derived for each subsystemic component; these can then be integrated within the model as a whole.

We envision that each hierarchical level can be adapted to utilise alternative estimation techniques depending upon the information available. For instance, the species richness of some taxonomic groups within some environments may be almost completely described, and a total species richness estimate (including undiscovered species) can then

be calculated utilising an appropriate statistical technique (Gotelli & Colwell, 2001). Other regions and taxonomic groups may use a similar, simple functional form such as equation (3). The framework that we propose is flexible enough to incorporate such disparate techniques and methodologies.

Once more turning our attention to species richness and endemism upon seamounts, the model outlined above can be incorporated within the decision-making process when the allocation of further sampling resources is under consideration. In particular, we would recommend that intensive sampling of a small number of seamounts is made a priority, in order that the species accumulation curve begins to level off in these regions, and asymptotic or parametric estimators (Gotelli & Colwell, 2001) can then be used to make predictions for the value of t_s and e_s . These values can then be used to further refine the functional forms in equation (3), and furthermore allow for a greater feasibility when incorporating hitherto ignored factors and parameters in an attempt to further represent the differences between seamounts. We would also recommend standardizing and collating the data that already exists – upon seamounts that vary widely in location, topography, and geophysical structure (Rogers, 1994) – in order that quantification of these differences can further proceed using both theoretical and empirical approaches.

References:

de Forges, B. R., Koslow, J. A., and Poore, G. C. B., 2000. *Diversity and endemism of the benthic seamount fauna in the southwest Pacific*. *Nature*, **405**: 944-947.

Gotelli, N. J. and Colwell, R. K., 2001. *Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness*. *Ecol. Letters*, **4**: 379-391.

Hilborn, R., and Mangel, M., 2003. *The ecological detective: confronting models with data*. Princeton University Press (Princeton, NJ).

Johnson, J. B. and Omland, K. S., 2004. *Model selection in ecology and evolution*. *Trends Ecol. Evol.*, **19**: 101-108.

MacArthur, R.H. and Wilson, E. O., 1967. *The theory of island biogeography*. Princeton University Press (Princeton, NJ).

May, R. M., 1988. *How many species are there on earth?* *Science*, **241**: 1441-

Neigel, J. E., 2003. *Species-area relationships and marine conservation*. *Ecol. Applic.*, **13**: 138-145.

Parin, N.V., 1991. *Fish fauna of the Nazca and Sala Y Gomez submarine ridges, the easternmost outpost of the indo-west Pacific zoogeographic region*. *Bull. Mar. Sci.*, **49**: 671-683.

Rogers, A. D., 1994. *The biology of seamounts*. Adv. Mar. Biol., **30**: 305-352.

Rohde, K., 1992. *Latitudinal gradients in species diversity: the search for the primary cause*. Oikos, **65**: 514-527.

Rosenzweig, M. L., 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge, U.K.

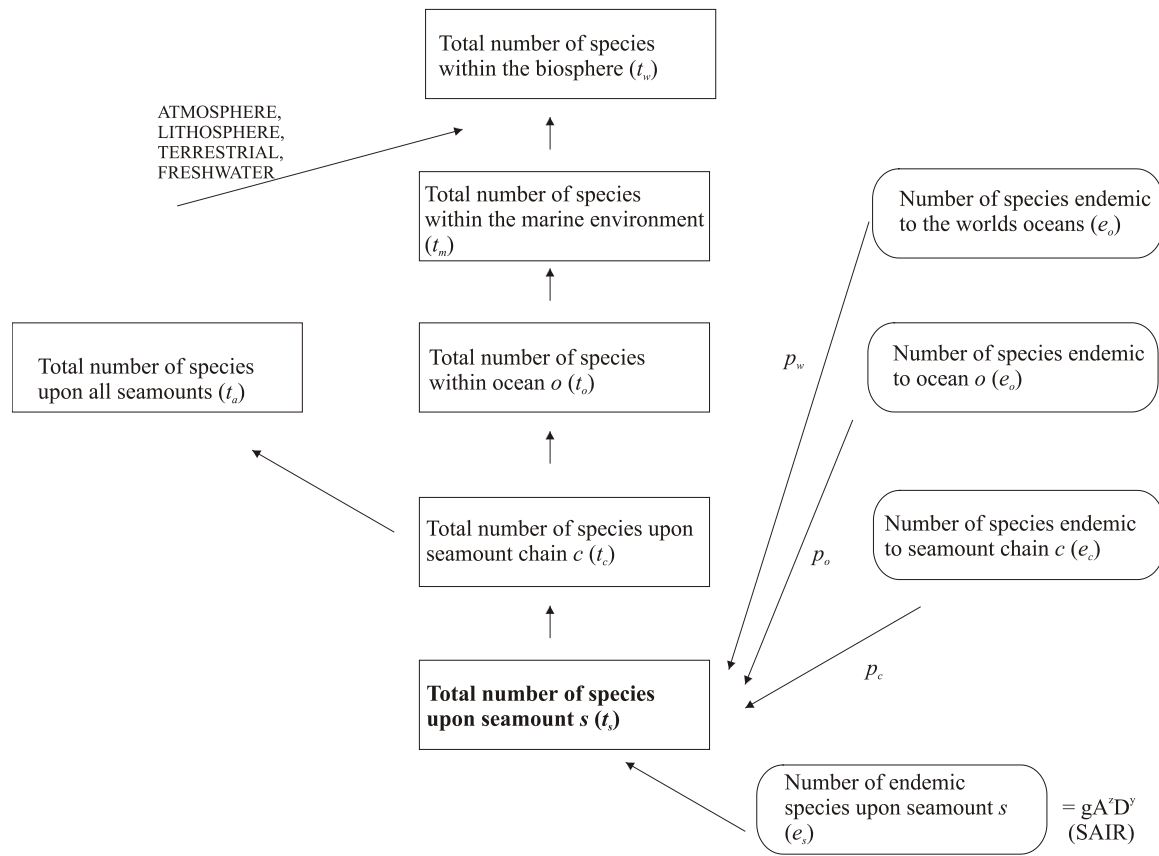
Appendix 1: Parameters

Parameter	Description
n	Number of seamounts
r	Number of seamount chains
u	Number of oceans
e_s	Number of species endemic to seamount s
e_c	Number of species endemic to seamount chain c
e_o	Number of species endemic to ocean o
e_w	Number of species endemic to more than one ocean or sea
A	Area of seamount s
D	Isolation distance of seamount s
x	Characteristic maximum dispersal range
c	Non-dimensional parameter in SAR
z	Non-dimensional parameter in SAR
b	Non-dimensional parameter in SAIR
y	Non-dimensional parameter in SAIR
t_s	Total number of species upon seamounts s
t_o	Total number of species within ocean o
t_m	Total number of species within the marine environment
t_h	Total number of species within the hydrosphere
t_t	Total number of terrestrial species
t_a	Total number of species within the atmosphere

t_l	Total number of species within the lithosphere
t_{dup}	Number of duplicate species present within two or more of the following environments: hydrosphere, lithosphere, atmosphere, terrestrial
$p_c(j)$	Binary function representing the presence/absence of the j th species endemic to seamount chain c upon seamount s
$p_o(k)$	Binary function representing the presence/absence of the k th species endemic to ocean o upon seamount s
$p_w(l)$	Binary function representing the presence/absence of the l th species endemic to the world's oceans upon seamount s
$q_o(l)$	Binary function representing the presence/absence of the l th species endemic to ocean o upon any seamount within that ocean
$q_w(m)$	Binary function representing the presence/absence of the m th species endemic to the world's oceans upon any seamount
$r_w(m)$	Binary function representing the presence/absence of the m th species endemic to more than one ocean/sea within ocean o .
c_s	Binary function representing the containment (1) or disjunction (0) of seamount s within ocean o (topographically speaking)
c_c	Binary function representing the containment (1) or disjunction (0) of seamount chain c within ocean o (topographically speaking)

Figure 1:

A SIMPLIFIED HIERARCHICAL SPECIES RICHNESS MODEL UTILISING ENDEMISM AS A BASIS



THINGS TO ADD

- (1) SUGGESTIONS THAT CAN BE DRAWN FROM THE SEAMOUNT MODEL IN TERMS OF SAMPLING – E.G. SAME NUMBER OF SPECIES FOR A LARGE SEAMOUNT VS. A CHAIN OF SMALL SEAMOUNTS?
- (2) A ‘RECOMMENDATIONS FOR FURTHER SAMPLING’ SECTION.
- (3) MINIMUM CUTOFF SIZE FOR SEAMOUNTS TO BE CONSIDERED NON-ISOLATING TO OTHER LOCALS.