

# Optimal Sampling for Complexity in Soil Ecosystems

**Arturo H. Ariño**

Department of Zoology and Ecology  
University of Navarra  
artarip@unav.es

**Carlos Belascoáin**

TB-Solutions S.A.  
belascoc@tb-solutions.com

**Rafael Jordana**

Department of Zoology and Ecology  
University of Navarra  
rjordana@unav.es

## 1. Introduction

### 1.1. Complexity in Soil Biology

Complexity in soil biology is a multi-level concept. Soil itself is the result of multiple interactions between physical structure, interface phenomena, soil biota activity, population dynamics, chemical composition, time, and environmental conditions. In turn, the resulting system (the soil) influences all those factors except time. Soil complexity can thus be observed at different physical levels (i.e., frequency distribution of aggregates' sizes, order of strata, etc.), biological levels (i.e., taxocoenoses, oxidable organic matter availability, population distribution, etc.), interaction levels (i.e. mineral paths between compartments, food web, etc.), or evolutionary levels (short-term variations on water availability, long-term erosion, etc.).

Dealing with all the complexity of soil structure at all these levels, and its evolution, may be beyond our current ability. However, we might try to capture some measure of at least the complexity of some of these levels. Should we be able to measure that at different levels, we could use these measures to construct a model that could relate the underlying processes from whence the complexity emerges. In doing so, it should be taken into account that simple models of complex ecological systems tend to be closer to the truth than more complex models of the same systems (Mikkelson, 2001).

Soil ecology studies whose main purpose is to serve another goal, i.e. community characterisation or biodiversity determination (which, in turn, may be used as tools for yet another purpose, i.e. "biotope quality" assessment, as indicators - Kurtz et al., 2001) can be regarded as a way to obtain a simplified view of a naturally complex system. Their goal is

often to produce a “black box” that ensnares a given set of complex factors, and then concentrate on the inputs and outputs of that subsystem as related to the rest of the system. How these relations affect the system often represents “a more useful kind of truth” (Mikkelsen, 2001) than the sheer number of factors that actuate within the black box.

Complexity in soil systems is, naturally, more than simply the number of these black boxes and their interactions. However, one can look for operators that can encode a great deal of information, yet convey clearly some of the essence of what makes the system complex (Paul K. Davis, personal communication). Biodiversity estimates can fit within this category, as they are related to how the soil has come to be as a result of the interactions depicted above. It has already been empirically demonstrated that there is a cross-scale influence of complexity levels when considering different aspects of the soil system, such as the close relationship between animal biodiversity and physical soil structure (Jordana et al., 2000) or even chemical characteristics (Peltier et al., 2001).

## **1.2. Soil Sampling Issues**

These soil ecology studies usually require processing a series of samples, that must meet the basic conditions for ecological studies: appropriateness, homogeneity, objectiveness, standardisation and efficiency (Gauch, 1982, p. 44-8). The fulfilment of these conditions would make data from several isochronal samples comparable (it is another matter to sample along time). Ideally, a soil fauna sample should be large enough so as to allow the characteristic(s) under study to be measured under no more error than acceptable statistical sampling error. However, when it comes to quantitatively determine the soil zoocoenosis (which is necessary for diversity or richness measurements), the spatial distribution of many soil fauna populations, and very frequently their aggregated nature, must also be taken into account. Thus, one may choose between taking a very large sample that encompasses a number of clumps, or taking a set of smaller, random and independent (Kasprzak, 1993) subsamples that are spread over a larger area and accumulate the information drawn from these.

A sufficiently large number of subsamples should statistically gather both the species richness information, or purely alpha diversity, and the beta diversity, that is, the portion of diversity that comes from the spatial distribution of the individuals of the populations (Margalef, 1980; Caswell and Cohen, 1993). It is well known that the number of species usually increases with the sample size, and so does the diversity (Margalef, 1980; Magurran, 1989; Huston, 1994; Roszenweig, 1995). Aggregating the results of several subsamples would thus yield a value of diversity usually larger than the value derived from one single subsample. If it is necessary to determine the diversity of the whole sample, then the size of each subsample, as well as the number of subsamples that make up one sample, have to be determined. Very small subsamples will result in a large variance between subsamples; but very large subsamples may mean an unnecessary amount of material to process.

In spite of the importance of ensuring a sufficiently representative sample, ecologists have not always appreciated the true effect of the sampling effort for the assessment of biodiversity (Gotelly and Colwell, 2001). Extracting and classifying soil fauna, which are necessary for determining diversity, is not an easy task. If the study deals with large areas

from which many samples are taken, i.e. transects across ecotones or heterogeneous regions, the number of samples to process may rapidly reach the “unmanageable” level. It is thus of paramount importance to determine which is the minimal sample size for these type of studies, that will yield representative results without imposing excessive (and unnecessary) work (Kasprzak, 1993).

The problem of minimal sampling has been constantly attacked over time. There are several “rules of the thumb”, as well as formal methods, regarding the increase of the number of species in the sample vs. the size of the sample, or number of subsamples for soil diversity studies. Optimisation of the sample size would be ideally achieved by sequential sampling (Krebs, 1989, p. 238) where only the strictly necessary number of subsamples would be taken; but this is usually not feasible for soil studies as the process of the subsamples is a lengthy one (Krebs, 1989, p.196) and may include even the determination, to species level, of the zoocoenosis. It is a common strategy, thus, that studies dealing with comparison of soil fauna diversity across several biotopes may need a systematic sampling where the minimal sample size is to be determined beforehand in a pilot study, as it is done in other, perhaps less complex systems where a single ecological indicator is sought (Kurtz et al., 2001).

### 1.3. Accumulation Curves

Most usual methods of prior minimal sample size determination would use some derivative of the species–accumulation curve. The sample size is plotted against the number of species, and minimal sample size correspond to one of several “stopping criterion” that in general mean that, beyond a certain sample size, the increase on species richness measured as species number is low enough so as to justify not taking a larger sample.

When “diversity” is meant to be just “richness”, i.e. number of species, there are a number of parametric and non-parametric estimators that will give the number of species of standardised, smaller samples (such as rarefaction) or extrapolate to standardised, larger samples (such as Michaelis-Menten curve or Chao estimator). Most of them rely on randomisation of existing data, such as collected individuals, to get the necessary variance estimates. A number of computer programs are routinely used for these estimations; the most popular ones being EstimateS (Colwell, 2000), WS2M (Turner et al., 2001), and EcoSim (Gotelli and Enstminger, 2001). It should be noted that these methods will yield and estimate of the total richness of the biotope sampled, much better than can be obtained from the direct reading of these parameters from the subsamples themselves; but they do not ensure that the asymptotic (total) richness has been reached (Gotelli and Colwell, 2001).

A minimal sample size for diversity studies of soil fauna has to ascertain that the measured diversity, and not just the species richness, is representative of the biotope being studied. Therefore, it is the accumulation of diversity what should be used in order to determine that critical size (Magurran, 1989, p. 59). A diversity-area curve would measure both alpha and beta diversity, and would ideally flatten at the point where the total diversity of the biotope is measured, i.e., the structure of the community under study is caught. (We do not take into account now the gamma diversity, *sensu* Margalef, or diversity through time).

However, to construct such a curve from the aggregation of results of smaller subsamples, as it would be the general case, or by adding specimens one by one, has also a major drawback. The order in which subsamples, or specimens, are added to the curve significantly influences the value of diversity reached at each point. The joint list of species and their abundances is different depending on which subsamples from a larger sample are added up. Thus, one may obtain different diversity curves depending upon the particular choice of subsamples and the order in which they are added to obtain the progressively larger area's species list. Therefore, the selection of minimal sample size by any of the usual subjective criteria is also affected, yielding different possibilities. This effect parallels the one observed when constructing species accumulation curves, and Monte Carlo randomisations of subsample aggregates are the choice methods for the construction of smooth species accumulation curves that can be treated analytically or statistically (Christen and Nakamura, 2000). Again, these methods may permit to infer the richness of the biotope being sampled from the bootstrap samples, but cannot ascertain that the diversity measured statistically correspond to that of a complete enumeration of the biocoenosis being sampled. This could be accomplished by somehow ensuring that the samples being analysed are true minimal samples of the biotope *for* the parameter being measured.

Accumulation curves can be used for this purpose. An asymptotic curve could, by definition, yield the maximum value of the parameter once the asymptote has been reached, and that would in turn give the value of the parameter and no estimate would be necessary. However, this is not usually the case, and a stopping criterion must be chosen.

Contrasting to the analytical and statistical treatment of inference from accumulation curves, the stopping criterion for an accumulation curve has been somewhat left to the researcher's discretion. Most scientific literature that resorts to a species-area curve (a species accumulation curve where the x-axis is not individuals but sampled area which prevents it from analytical description and requires bootstrap sampling and randomisations –Gotelli and Colwell, 2001-) for minimal sampling assessment do not explicitly develop a formal stopping criterion other than “seeing the flattening” of the accumulation curve or other arbitrary criteria.

## **2. Minimal Sample Size for Diversity Measurement**

We propose a method for determining the minimal sample size by calculating the accumulation of diversity from standardised subsamples, that deals with the problems explained above. Our method is based on the accumulation curve for the parameter of interest, which can be a measure of the complexity of the system. As a stopping criterion, we resort to the slope analysis, a family of methods used for two-sample comparisons (Dawson, 1998). We construct a diversity accumulation curve from subsamples, where each point represents the average diversity of all possible combinations of subsamples, or of a bootstrap sample of these combinations. This allows us to determine the point of non-significant diversity variation, and thus minimal sample size, by a simple statistical test on the slopes of the accumulation curve points. The method can be used to assess the number of subsamples that will be necessary to process in order to estimate diversity from soil fauna samples on large studies.

## 2.1. Method

Let it be  $N$  subsamples of equal dimensions (be them surface, volume, or mass) coming from a pilot study. We intend to determine the  $M$  number of subsamples that conform the minimal sample for calculating the total diversity. Each subsample is processed, and their faunistic contents are extracted and classified. The total extracted faunal categories (taxa) are  $T$ . The abundance results can be represented as the usual  $NT$  array, where columns are subsamples and rows are the taxa.

One will first calculate the chosen diversity parameter  $D$  of each subsample, which is assumed to be subminimal. We thus obtain a series of  $N$  individual values of diversity. Next, we combine the first and the second subsample, and obtain the diversity of this new double-sized, pooled sample. We do the same for all possible pairs of subsamples, that is, all possible combinations of  $N$  vectors taken by twos, obtaining a new series of diversity values for all possible pairs. The algorithm thus continues by combining all possible trios, quartets, and so forth, of the subsamples and always calculating a set of diversity values for each series. The number of possible combinations of  $x$  subsamples within the sample drawn from the  $N$  total samples,  $N!/x!(N-x)!$ , can be a rather large figure; in that case, the algorithm will select a random set of those combinations.

In the next step, all series of diversities for the combinations of 1, 2, 3 ...  $x$  ...  $N$  subsamples are averaged. We thus obtain a single value of average diversity, along with its variance, for each sample size. Plotting the average diversity for each sample size against the sample size, which is the number of subsamples  $x$ , we obtain the familiar diversity/area curve.

Now we select the stopping criterion. For the purpose of measuring complexity in the form of diversity, we assume that we have captured the diversity of the system when additional samples do not add additional value to the diversity index that we have chosen, that is, when the slope of the curve between a pair of consecutive points is not significantly different from zero (i.e., the curve is statistically "flat"). Thus, we can perform a simple  $t$ -test for the mean of the slopes between each averaged diversity measure and all possible values of the following one, and stop when we cannot reject the null hypothesis of these slopes averaging zero at the desired significance level.

## 2.2. Algorithm Implementation

The algorithm can be summarised as textual metacode in the following steps:

1. Read the data file, which will contain records in a suitable format that will express the abundance of each taxon in each subsample.
2. Create a vector  $\mathbf{S}[1...N]$  containing the subsample codes.
3. Obtain  $N$  vectors  $\mathbf{X}[1... \binom{N}{x}]$  of all possible combinations of  $N$  subsample codes from  $\mathbf{S}$  taken by  $x$  where  $x (1 \bullet N)$ .
4. For each of the  $N$  vectors  $\mathbf{X}$  do:

- 4.1. Create a new empty vector  $\mathbf{D}_n[1 \dots \binom{N}{n}]$  for the  $D$  values.
  - 4.2. For each element in  $\mathbf{X}$  do:
    - 4.2.1. Merge the faunistic lists of each subsample contained in the element, adding the abundances for equal taxa.
    - 4.2.2. Add a new  $\bar{D}_n$  element to  $\mathbf{D}_n$  with the  $D$  value obtained.
  - 4.3. Calculate  $\bar{D}$  for  $\mathbf{D}_n$ .
  - 4.4. Compute one-sided  $t$ -test for  $H_0=0$  for the slopes between  $\mathbf{D}_n$  and  $\bar{D}_{n-1}$  if it exists, against  $H_1$ =greater mean.
  - 4.5. Plot  $\bar{D}$  against  $n$ .
5. Choose sample size= $n-1$  when  $H_0$ =true at the desired significance level.

### 3. Discussion

While this diversity accumulation algorithm does not avoid initial oversampling, for it is necessary to combine several small subsamples in order to be able to statistically compare the means of their diversities, it may save extra work if used as a pilot study prior to larger, systematic studies by assuring that the samples taken will have adequate size. Whereas this is the goal of any sampling optimisation technique, we provide a way to do it for repetitive diversity measurements that includes some “statistical certainty” that, on average, we are dealing with samples that will represent the true (up to beta) diversity of the biotope.

Our choice of averaging the slopes between the average diversity of all combinations of subsamples and all the possible combinations of subsamples at the next point may overshadow another possible technique, perhaps more straightforward, that would consist on drawing all possible diversity accumulation curves for a set of subsamples, calculating the average slope of all curves at each point and stopping when the average slope does not significantly differ from zero. Though this may be more intuitive, we discarded it for the sake of parsimony: There are  $N!$  possible accumulation curves for a given set of  $N$  subsamples and  $2^N-1$  possible diversity values giving

$$\sum_{i=1}^{N-1} \frac{N!}{i!}$$

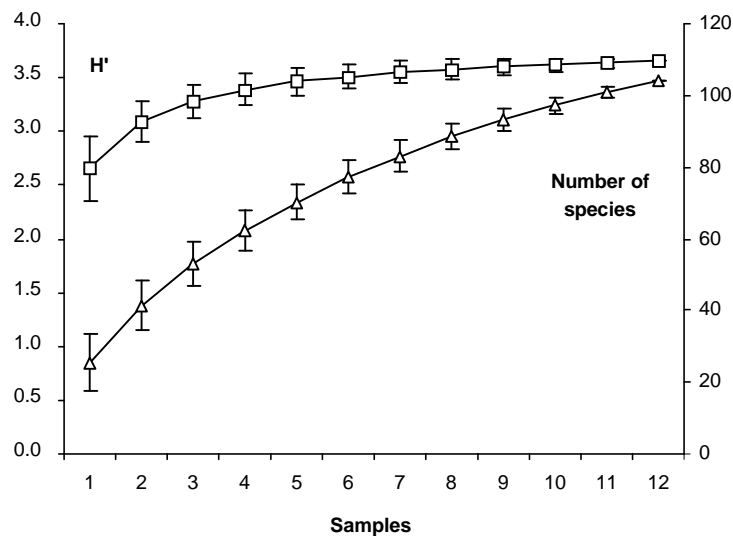
possible slopes between these values along the curve. However, the number of possible slopes if only the slopes between the average diversity at one point and all possible diversities at the next point are used is

$$\sum_{i=2}^{N-1} \binom{N}{i}$$

which is a much smaller figure.

The algorithm can be regarded as a null model that deliberately excludes the mechanism of diversity formation (Gotelli, 2001), as it seeks a stopping criterion that is explicitly statistical in nature and that does not depend on what the parameter means. Thus, any complexity measure could, in principle, be subject to the same algorithmic decision if they yield an accumulation curve.

As null models are usually based on real data sets of an index of community structure (Gotelli, 2001), we have used a set of 12 initial soil nematode subsamples taken from an oak forest in Navarra, North of Spain, as a part of a project on forest soil recovery after fire. Samples were taken by a cylindrical corer, 25 mm in diameter. The original TxN matrix, species level, is 104 x 12. The algorithm was programmed as a set of pure C routines, debugged and tested with a series of benchmark files, and the program was applied to the real data matrix. The accumulation curve for species for Shannon's diversity  $H'$  (fig. 1) shows no distinct flattening, whereas the slope of the diversity curve becomes not significantly different from zero (95% confidence level) after the tenth sample. The total number of  $H'$  values calculated were 4,095. The number of necessary slopes to test if all individual curves had been plotted to look for non-significant slope change would have been over 823 million.



**Figure 1.** Species/area and  $H'$  diversity/area curves when combining samples as described in the text. Values are averages, rounded to integer in the case of species. Error bars represent standard deviations. The mean of the slopes between the mean diversity at point 10 and the diversity values at point 11 become not significantly ( $p < 0.05$ ) different from zero, whereas the species numbers are always significantly different from each other.

## 4. Acknowledgements

Samples used in the example provided were taken by the Department of Zoology and Ecology of the University of Navarra during the Project ‘Study of the impact of forest fires and their recovery: soil science, fauna, flora, and erosion’, project no. FOR 90-0982-CO2 of the National Institute of Agronomic Research, Spain. C. Belascoáin was successively a grantee of the Friends of the University of Navarra Association and of the above Institute.

## 5. References

- Caswell, H., & Cohen, J.E., 1993. Local and Regional Regulation of Species-Area Relations - A Patch-Occupancy Model. In: R.E. Ricklefs and D. Schluter (Editors), *Species Diversity in Ecological Communities: historical and geographical perspectives*. University of Chicago Press, Chicago, pp 99-107.
- Christen, J.A., & Nakamura, M., 2000. On the Analysis of Accumulation Curves. *Biometrics*, **56**, 748-754.
- Colwell, R.K., 2000. *EstimateS: Statistical Estimation of Species Richness and Shared Species from Samples (Software and User's Guide)*, Version 6. <http://viceroy.eeb.uconn.edu/estimates>
- Dawson, J.D., 1998. Sample Size Calculations Based on Slopes and Other Summary Statistics. *Biometrics*, **54**: 323-330.
- Gauch, Jr., H.G., 1982. *Multivariate Analysis in Community Ecology*. Cambridge University Press, Cambridge.
- Gotelli, N.J., 2001. Research frontiers in null model analysis. *Global Ecology & Biogeography*, **10**, 337-343.
- Gotelli, N.J. & Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379-391
- Gotelli, N.J. & Entsminger, G.L., 2001. *EcoSim: Null models software for ecology*. Version 7.0. Acquired Intelligence Inc. & Kelsey-Bear. <http://homepages.together.net/~gentsmin/ecosim.htm>.
- Jordana R, Arpin P., Vinciguerra M.T., Gonzalez S., Aramburu M.P., Ariño A.H., Armendariz I., Belascoain C., Cifuentes P., Clausi M., Escribano R., Garcia Abril A., Garcia-Mina J.M., Hernandez M., Imaz A., Moraza M.L., Ponge J.F., Puig J. & Ramos A., 2000. Biodiversity across ecotones in desertifiable Mediterranean areas. In: Balabanis P., Peter D., Ghazi A., Tsogas M. (Eds.): *Mediterranean Desertification Research results and policy implications*, Vol 2.: pp 497-505. European Commission EUR 19303.
- Huston, M.A., 1994. *Biological Diversity. The coexistence of species on changing landscapes*. Cambridge University Press.
- Kasprzak, K., 1993. Selected aspects of mathematical statistics. In: M. Górny and L. Grüm (Editors), *Methods in Soil Zoology*. Elsevier, Amsterdam, pp. 16-69.
- Krebs, C.J., 1989. *Ecological Methodology*. Harper & Row, New York.
- Kurtz, J.C., Jackson, L.E. & Fisher, W.S., 2001. Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development. *Ecological Indicators*, **1**: 49-60.
- Magurran, A., 1989. *Diversidad ecológica y su medición*. Ediciones Vedral, Barcelona.
- Mikkelsen, G.M., 2001. Complexity and Verisimilitude: Realism for Ecology. *Biology and Philosophy*, **16**: 533-546.
- Margalef, R., 1980. *Ecología*. Omega, Barcelona.

- Peltier A., Ponge J.F., Jordana R. & Ariño A., 2001. Humus forms in Mediterranean scrublands with aleppo pine. *Soil Science Society of America Journal*, **65** (3): 884-896
- Roszenweig, M.L., 1995. *Species diversity in space and time*. Cambridge University Press.
- Turner, W., Leitner, W. & Rosenzweig, M., 2001. WS2M. Software for the measurement and analysis of species diversity (Software and User's Manual). <http://eebweb.arizona.edu/diversity>