Prospects and expectations: reflections on a science in change

L. Orlóci

Department of Plant Sciences, The University of Western Ontario, London, Canada N6A 5B7,
E-mail: lorloci@uwo.ca, Web: http://sites.netscape.net/lorloci

Abstract: I use "prospects" to connote the continuing subversion of the environment. "Expectations" refer to what I think Statistics should be making of itself in the era of global change science. I develop thoughts on these in three main sections. In the first, I recount up-to-date information about global warming. This topic is intriguing to me, since global warming is potentially the most subversive environmental process, appears least preventable, and as such, most deserving to be a focal point in ecological study scenarios. In the second section, I discuss briefly the ecological agenda and the implications therein for definition of the sampling environment. The third and last large section explains how I think of what should constitute an applicable statistical dialect for users in global change science. I detail the basic principles, enumerate choices, and deal with other bare necessities upon which such a Statistics should be founded.

General introduction

I decided to begin the Introduction with a short story, knowing well that it may not strike the reader as being quite "scientific" in the manner of the usual papers published by Community Ecology. My reason is to recapitulate at least in a limited way the mood of the original oral presentation[1]:

> The story is about the flamboyant confederate cavalry general, Nathan Bedford Forrest, about whom it is said to have had 30 horses shut out from under him, but he killed 31 union cavalry, and in the balance, he ended up with a one-horse advantage. The story takes place in December 1862 at Parker's Cross Roads in Tennessee. Forrest is commanding the raid on the union force assembled in the area, but stops fighting on white flag. He is fixing to take enemy prisoners when his staff officer calls out. General! Union cavalry to the rear! What shall we do? What shall we do? Divide and charge both ways!! — came the answer. The charge resumes, enlarged to go forward and to back, while Forrest is manoeuvring to withdraw sideway. Forrest's tactics worked. He and his troops outmanoeuvred a superior enemy force. The morals? Tactics and gallant men can win battles. But only when the action is part of an effective global strategy can the battles won build toward winning the war. The confederates lost the war.

Ecology has become front-line science of global change. The ecological approach is built about a series of short-term tactical and long-term strategic objectives. The tactics are more likely to be under the ecologist's control, since these address problems that are, in large part, local to the conduct of good science. The strategic problems, such as the upholding of global environmental quality and high biodiversity, are different. Their dimensions span science and politics, and as in such, strategic action is slow and the conclusion is a protracted process. Why is it so? For one thing the depth of the problem recognised has to be assessed and appropriate conceptual and technical tools have to be developed to further the effort. For another thing, the public and political leadership have to be kept informed at their pace and plans have to be accepted. The latter engages the ebbs and tides of the political game in which the controls are not in the ecologist's hands.

Things being that way, questions about what should we regard as high-level strategic problems and what should be our contributions to the solutions are intriguing. My first thought on these comes from common sense: by the time the full-blown problem is seen to be as such, it may already be too late to do much about it. Foreseeing

problems in the making, and preventing them from becoming unmanageable monsters, would be a sensible way to go. Global warming is an éclat case in point. It is true that a good measure of ambivalence still exists about the reality of global warming, but the ambivalence is eroding. Respected authorities on the subject — most recently in reports by IPCC[2] and NSF[3], the latter commissioned by the White House — came to the conclusion that global warming is in fact an ongoing process. There is in it the potential for enormous destruction which should immediately elevate global warming into the category of the "most urgent strategic problems".

The danger of global warming notwithstanding, the apparent remoteness of the problem, and particularly the relatively slow onset of definite manifestations, tend to lessen the sense of urgency for taking immediate preventive steps. These are typically factors in the U.S. government's decision to renounce the Kyoto accord (see Time's report, April 9, 2001, page 32). I would not imply from this that the U.S. government sees the potential dangers subsiding. Rather, I should take the decision as one echoing the industrial views:

- Global warming cannot be stopped.

- The limited positive effects to be gained by implementation of Kyoto are out of proportion with the foreseeable costs to industry.

- It is cheaper to pay for the damage when it already occurred than to start paying for Kyoto now.

The question is, of course, not if a next hard decision has to be taken but when should it be taken to help stabilise atmospheric CO_2 concentration. Hoping for the process to right itself, as advocated by B. Lonborg (2001)[4], cannot be considered an assurance that the process will not be enormously destructive. Furthermore, payment after the fact does not strike me as a particularly valid option, considering the distinct possibility of no economy will be left in a shape by that time to pick up the tab.[5]

I start with an abbreviated discussion of global warming and its effects on vegetation dynamics, in the main text. This is my starting point in considerations of the ecological agenda, its consequences in the definition of the sampling environment, and the consequences for Statistics direction. It is quite clear to me that any new contribution to contents has to be, in large part, in the area of new conceptual tools for users of statistical techniques. The new tools should entail all major lines of concern, tactical and strategic, such as sampling, inference, prediction, decision-making, policy design, etc. I develop thoughts on sampling, inference, and prediction with a view to the specialised aims and the uniqueness of the sampling environment of my field.

Prospects of global warming

I should go directly to the 2001 report of IPCC and consider its implications. The report is concerned with facts and projections, regarding global warming. The ambivalence that surrounded global warming not so long ago is quite clear from J. Mason's lecture on global warming delivered in 1990 at the Royal Institute under the sponsorship of British Coal:

> "... the current [1990] best estimates of global warming are not so alarming to warrant major strategic changes in energy, agriculture, etc."

It is true too that in the interim decade much new evidence came to light, mainly from new measurements and improved climate models, and the expressions of ambivalence are more and more muted. The evidence seen proved to be sufficiently compelling to IPCC to conclude in its 2001 report :

> "The globally averaged surface temperature is projected to increase by 1.4 to 5.8 $^{\circ}$C over the period 1990 to 2100." [6]

The projected 110-year range is equivalent to warming by 1.3 to 5.3° on a 100-year basis. This indeed implies "mega" warming, roughly 22 to 93 times the historic warming rate (0.0625°/100 yr) experienced by the Earth's

---

2   Intergovernmental Panel on Climate Change. 2001. Third Assessment Report. - Climate Change. Find report at web address: http://www.ipcc.ch

3   National Research Council. 2001. Climate Change Science: An Analysis of Some Key Questions. Washington, USA. Find abstract in "Global Change Digest: May/June 2001 Edition" on the web: hyperlink "http://www.globalchange.org/gccd/ Digest/homepage.html.

4   B. Lonborg's article appeared in The Economist, August 4-10, pp. 63-65.

5   Seeing the difficulties of the present political systems with decisive action just to stabilise CO_2 in the atmosphere, the ultimate long-term solution advocated by Thomas Berry (1990), namely adoption of an ecocentric mode of existence, is unfortunately receding farther and farther into the distant future. The long-term prospects over millennia are of course as clear as it is certain that the sun cycles continue and the Milankovitch cycles hold their own (Milankovitch 1941): the climate will cool, and with it, a new glacial period will return, give another few millennia.

6   The range in rates corresponds to differences in model scenarios. Uncertainties are involved, owing in good part to the speculative manner in which cloud properties and their interaction with the radiation fields are presented, and to the mode in which the models handle internal feedback (see Mason 1990, IPCC 2001, pp. 49, 66, 67).

surface on average during eight millennia that directly preceded the Hypsothermal 6000 years ago.

What were the effects of global warming on the Earth's biota when it came at the long-term $0.0625^o/100$ yr rate? For one thing, it was during those few pre-Hypsothermal millennia that the global vegetation pattern has attained its modern global map through a series of dramatic shifts. All regions of the Earth were affected, but most definitely the cold and temperate zones at higher northern latitudes. Delcourt and Delcourt (1987) present a good example of the shifts (their Figure 1.4, pp. 20-21) that culminated in the present continental pattern of vegetation formations in eastern North America (Figure 1). The regional dynamics of a typical case is pictured schematically in Figure 2.

It is quite clear how effective even a moderate warming process can be in generating very remarkable levels of vegetation dynamics. An example of this is the rise of Lucy Braun's Eastern Deciduous Forest (Braun 1950) as a distinct formation. This formation did not exist until about 14000 years ago. Its constituent species were sheltered in fragmented southern refugia. But then the climate started to warm, and the formation started to assemble and expand with North/South span exceeding 15 latitude degrees (D in Figure 1) at its maximum extent by the Hypsothermal. The statistics of expansion are posted in Figure 2.

Unexpected by many, global warming and cooling cycles of the Late Quaternary, which so dramatically affected the vegetation pattern in cold and temperate zones, had measurable effects also in the tropics. Figure 3 portrays a case from the Amazon of Brazil (see information in Table 1). The graphs show mirroring of diversity and velocity, and in a related sense, diversity and stability under climatic effects. It is seen that during the period cooling tended to promote vegetation stability while warming tended to undermine stability.

Under the $0.0625^o/100$ yr warming rate, plasticity and migration ability of the populations were still able to sustain temporal and geographic contiguity in the formations. Could something similar be expected to occur if global warming reached velocities at least 22 times the historic rate? I do not believe it could be. Warming at those velocities is bound to deteriorate the environment sufficiently within the life span of an average forest tree to cause stand-level dieback over entire forest regions in the cold and temperate zones (Orlóci 1994). The alpine arctic floras of tropical mountains could be entirely lost in sites like the Mauna Loa, Mauna Kea, and Haleakala in the Hawaiian Islands (Orlóci 1994). Under extreme cir-
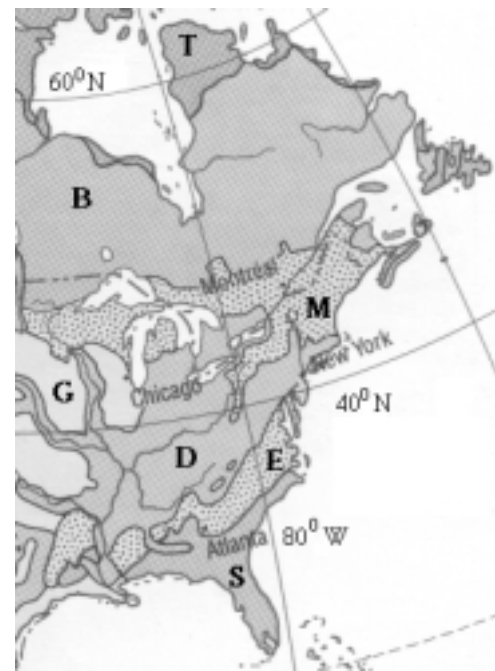

Figure 1. Vegetation map of Eastern North America. Legend: T – Tundra, B – Taiga/ Boreal Forest, M – Cool Mixed Conifer-Deciduous Forest, G – Grassland, D –Cool Temperate Deciduous Forest, E – Warm Mixed Evergreen Forest, S – Subtropical Forest. After the Rand McNally Classroom Atlas (1988) modified.
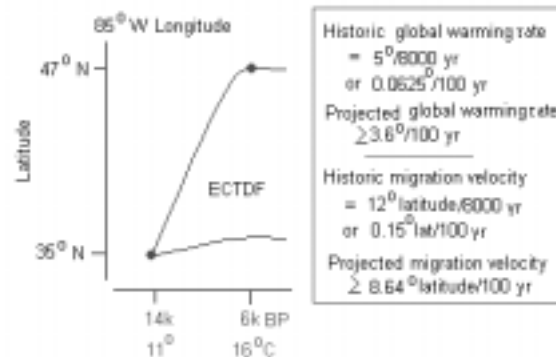
Figure 2. Migration dynamics of the Eastern Cool Temperate Deciduous Forest of North America under global climate warming during an 8000-year period of the Late Quaternary. Projected warming follows Manabe at al. (1990) with rate expressed on a 100 yr basis. Projected migration rate accords with $3.6^oC \times 0.15^oLat/0.0625^o$. Graph based on original in Delcourt and Delcourt (1987), idealised.

cumstances, the endangerment of global biodiversity that Juhász-Nagy (1994) portrayed so well is not expected to subside. In fact, the proposition that half of all the species alive today — plants, animals, lower and higher forms —

Table 1. The P. E. De Oliviera paleopollen data set. Data are posted on the world wide web at address hyperlink "http://www.ngdc.noaa.gov/cgi-bin/paleo/ftpsearch.cgi" # 19 Dec 97, # Percentages for the top 15 pollen types, # Site name: Lagoa das Patas, # Place: BRA:Amazonas, # Latitude: 0.16.00N (.266667), # Longitude: 66.41.00W (-66.683333), # Elevation(m): 300, # Contact person: De Oliveira, P.E., # Reliable age bounds: 0 – 42210, # 17 taxa, 49 paleorelevés in paleospectrum. Dark-shade: original scale. Light-shade: scale exaggerated for better viewing of trace taxa. Solid line across in body of graph: a paleorelevé. Letters across top: a — *Alchornea/Aparisthmium*; b – Aquifoliaceae; c — Arecaceae (excluding *Mauritia*); d — Caesalpiniaceae undiff.; e – *Cassia*; f – *Cecropia*; g — Combretaceae/Melastomataceae undiff.; h — *Copaifera*; e – *Mauritia*; j – Myrtaceae; k – *Podocarpus*; l – *Tapirira*; m — Urticaceae/Moraceae; n – Fabaceae; o – Poaceae; p — Other trees and shrubs; r — Other herbs.
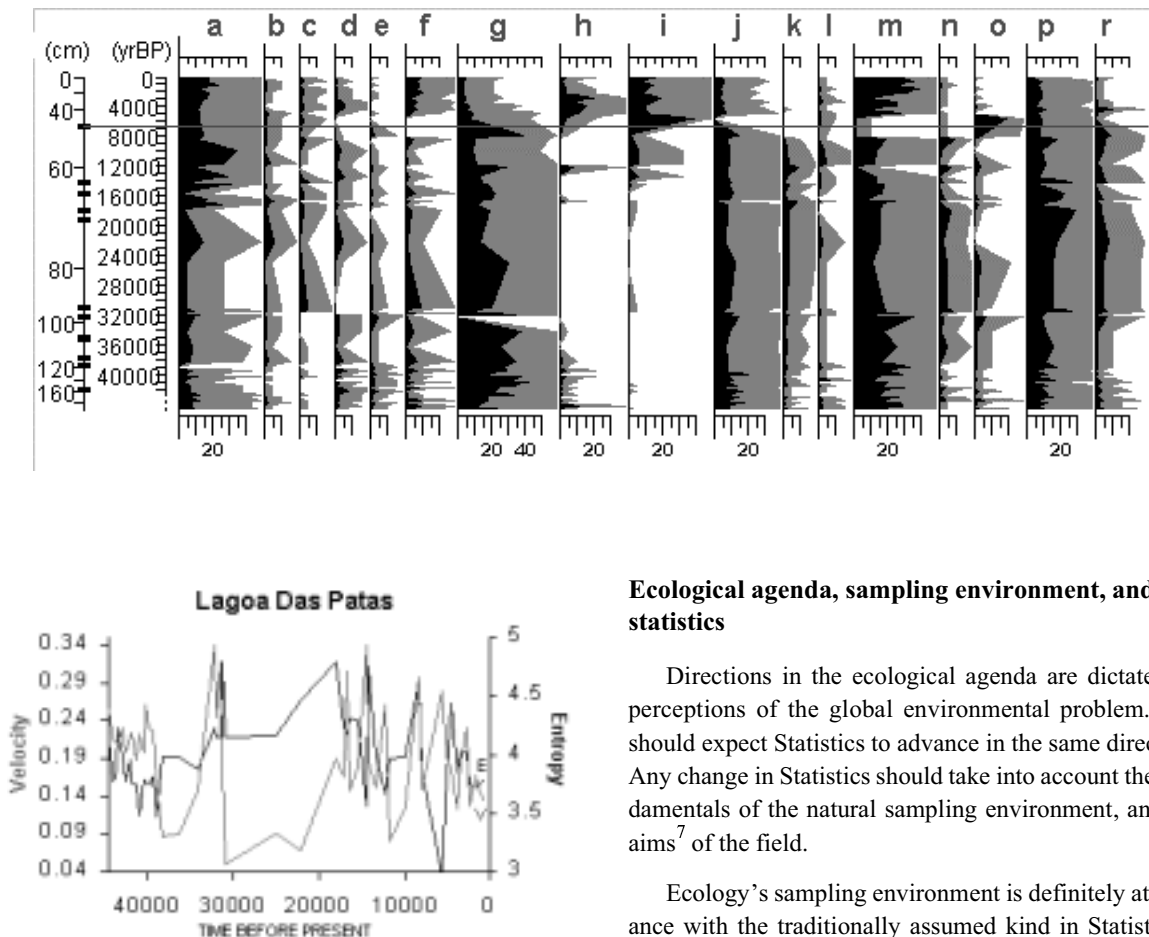
Figure 3. The evolution of diversity (entropy) and process velocity as reflected in the paleopollen spectra of Lagoa das Patas in the Amazon of Brazil. Information on the data set is presented in the caption of Table 1. See explanations in the text and in Orlóci (2000). Note that graphs start at second data point.

could be lost by the end of the 21st Century (see Lean and Hinrichsen 1992, pp. 127-129) may indeed prove to be a conservative projection.

Ecological agenda, sampling environment, and statistics

Directions in the ecological agenda are dictated by perceptions of the global environmental problem. One should expect Statistics to advance in the same direction. Any change in Statistics should take into account the fundamentals of the natural sampling environment, and the aims[7] of the field.

Ecology's sampling environment is definitely at variance with the traditionally assumed kind in Statistics. I refer to the conflicting conceptualisations in this regard as the *traditional ecological view*, an adherent of the ecological theory, and the *traditional statistical view*, an adherent of the classical small sample theory of statistics.

The traditional statistical view is closely entwined with the mentality that holds "normal" the anticipation of sampling environment homogeneity and response-interaction linearity. It also holds as being "normal", the presence of crisply defined sampling units, and an unhindered access to all units. With all these incredibly assumed, statistical sampling becomes an exercise in random selection

---

7   I should mention environmental assessment and biodiversity monitoring, and refer to two remarkable sets of documents. One is the superbly constructed Hungarian scheme for monitoring, in 10 volumes under the editorship of Horváth, Korsós, Láng, and Matskási (1997). The other is the *Encyclopaedia of Biodiversity,* recently published in 5-volumes under the editorship of S. A. Levin (2001).

of units, statistical analysis becomes largely a matter of tracing the behaviour of moments and product moments, and statistical testing becomes a matter of probability scaling based on theoretical distributions derivable from first principles, mostly from the Normal.

The traditional ecological view presents a fundamentally different kind of sampling environment. In this, heterogeneity rules, responses and interactions are likely to be non-linear, and the units fuzzy, neither "catalogued" nor with "unhindered" access. Dynamics is as given, and pattern, other than random, is a rule (Orlóci 2000, 2001a). The statistical dialect that can function in such an environment is to be called "applicable Statistics".

An applicable statistical dialect – general principles, bare necessities

Considering the ecological sampling environment, any new development in the line of the statistical approach or conceptual tools should obey the general principles of *flexibility, empiricism*, and *local relevance*. Since these are self-evident for the ecologist, I keep the comments rather brief:

1. *Flexibility.* A flexible approach has the characteristics of Poorean successive approximation (Poore 1962, Wildi and Orlóci 1991, Orlóci 1993). In this, the next-step very much depends on what has been discovered in previous steps. Refinements and reasoned changes in technique and direction are permitted at any step. Therefore, the conclusion condenses around itself within increasingly smaller radius. When it cannot be further condensed with the expenditure of reasonable effort, the conclusion becomes an inference.

2. *Empiricism.* The tenets of empiricism imply derivation of basic propositions about Nature in the course of observation and experiment. This is quite different from the approach that operates on the basis of the principle that conceptual convenience justifies replacing Nature with *a priori* surrogate mental constructs, like in the practice of conventional Statistics.

Reasoning from known principles is, of course, a strong component of the ecological approach. It can fill gaps, left blank in one's knowledge by things intangible in Nature whose existence is suspected from conditions found existing in other things. This is like observing an apple falling down and not up, hence the conclusion of the Newtonian force, gravity; or populations not staying put, but migrating to avoid destruction, hence the Cainean principle of organic evolution being slower than environmental change (Cain 1944).

3. *Local relevance.* When a conclusion is reached by means that rest upon unmitigated logical connections to the actual conditions of the sampling environment, local relevance is established. Generality is an opposite of local relevance only if it comes about in the presence of unreasonable assumptions that replaced actual conditions.

Clearly, the conceptual problems facing the practitioner in ecology are not some odd peculiarities, but rather they are part and parcel of normal ecological practice. So, an operational Statistics has to have solutions to real problems, rather than offer solutions to problems under sanitised surrogate conditions. Some questions:

1. *Sampling: random or preferential?* We sample because we cannot completely enumerate things that populate the sampling environment. Unfortunately, random selection is not, as a rule, a rational option (Pielou 1975, Edgington 1987, Orlóci 1993); the preferential selection of natural units is all but preordained. The result of this is a collection composed of *type specimens, typical sola, typical tree specimens, typical community stand*, and the likes. If random sampling were forced, it would surely violate the principle of frugality, and likely, also the principle of representativeness. The former implies collection of much useless materials, such as incomplete plant specimens, spatially heterogeneous vegetation units, and so forth. After having said these things, let me try to comfort (hopefully not to further distress) the statistician in us by drawing attention to one strongly held belief in ecology:

Much of what is known about the biota on all scales from life zone to molecule, and about the environment on all scales from Ice Age climates to pedogenic processes, is the consequence of smart people looking closely at materials that they considered suitable for analysis; rather than bothering with materials that chance handed to them in blindly executed sampling.

2. *Responses, interactions: linear or non-linear?* The point should not be missed: in the ecologist's world nonlinearity is the rule (Figure 4) and linearity is exception (Orlóci 1980, Podani 2000). I have seen non-linear response curves, like those in Figure 4, described as the Groenewoud-Whittaker type, or Gaussian type, but the reader should keep in mind that these refer to curve shape, not to a frequency distribution. As pictured in Figure 4, two species populations (X_1, X_2) are responding to influences along a major environmental gradient. The gradient is reasonably long, and as expected, the populations' joint scatter is a horseshoe (Figure 5). The consequences of having a horseshoe-like joint scatter, or something even more complex, like a spiral in higher dimensions, can have devastating effects on the reliability of traditional
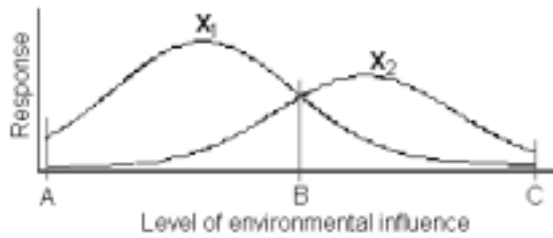
Figure 4. Groenewoud-Whittaker type species responses to environmental influence on a long gradient. Idealised mean potential responses are pictured. In reality, the curves are the central axes of curved point clusters.
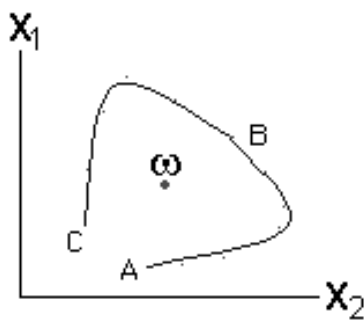

Figure 5. Central axis through a horseshoe point cluster. The centroid is at point ω.

statistical manipulations. The reason is, in part, their use of moments/product moments and the failure of the Normal density function

$$f(X_1, X_2, ..., X_p) = Be^{0.5(\mathbf{X}-\mu)'\Sigma^{-1}(\mathbf{X}-\mu)}$$

to be a meaningful construct (see Orlóci 1993).

3. *In focus: centroid or type*? When the sample consists of typical units, the behaviour of the "type" comes into focus. Does the behaviour of the centroid (centre of gravity, arithmetic mean vector as in Figure 5) still have utility for us? The centroid could replace the "type" if the responses were tested and found to be linear (Orlóci 1993). Should an actual case fall outside this specific type of response, such as for instance in the ecologist's horseshoe configuration (Figure 5), the centroid will likely reside within the void (ω) where no natural state
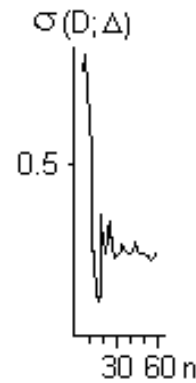
Figure 6. Graph of stress function σ(**D**,Δ). Two distance configurations (**D** and Δ) are compared, both *nxn* symmetric matrices. **D** depicts the vegetation configuration, and Δ the environmental configuration. Note the sharply declining value of stress with *n* increasing to about 20, and the stress oscillations in a narrow band about stress value 0.25.

materialised. In such a case, the centroid is the least likely state and it cannot represent the "type".

4. *Stopping rule in sampling: error variance or structure stability*? Stopping rules and reliability measures control the sampling process. Two properties are particularly relevant: sampling variance and sample structure stability (Orlóci and Pillar 1989). The sampling variance is measurable by $SV = V/n\,[1\text{-}n/N]$, written in the manner of Sampford's (1962) text. When a stopping rule is built around SV, the sampling process will stop if the sample size n is reached, preordained by the arbitrary choice of values for variance V and population size N. Clearly, the larger is n, the smaller the value of SV. So, there is no unique n to be found under this rule, except $n = 0$ or $n = N$, but these are utterly trivial. Ecological studies will do better when the stopping rule is formulated about the sample structure stability. What does this involve? Examination of Figure 6 should help clarify the point. Two structural descriptions of the same n sampling units are given. The structures are distance configurations, one $\mathbf{D}_{nxn}$, calculated from vegetation data, and the other Δ_{nxn}, calculated from environmental data. Optimal sample size is reached when the graph of σ(**D**; Δ) starts levelling off.[8]

5. *Incorporation of the time/space scale*. Perception is affected by scale. Therefore to see Nature enfold in its full glory all propositions have to be examined through scales. Conclusions regarding the principal forcing factor in vegetation dynamics represent an éclat example. Some

---

8   The criterion σ(**D**; Δ) is a measure of stress or discordance. The zero point on its scale is always unique, but the upper limit can be case dependent. If it is, the nominal σ(**D**; Δ) is not an appropriate guide to measure the relative strength of linkage between **D** and Δ. It is important though, the actual magnitude of σ(**D**; Δ) does not affect determination of the locally optimal sample size.

may think it is the climate, others may point to facilitation, while still others prefer to focus on dominance sorting by population processes (Orlóci 2000). In fact, all things act as parts of one system, and none alone by itself. The question is: which one should be regarded most important? The answer will differ depending on the scale at which the process is examined: climate for formation level patterns, facilitation for zonal/azonal community patterns within the formations, and dominance sorting for patch level arrangements or in gap dynamics.

6. *Unconventional phase space properties in focus*. The statistical practice of concentrating on the behaviour of moments and product moments is quite arbitrary. The practice can even be harmful to the users reputation when non-linear structures are present. In all cases there should be much utility gained by focussing on the statistical behaviour of other physical properties, such as *shape, form, and chance* (fractal, Mandelbrot 1977, Scheuring 1993, Walker and Kenkel 1998), *degree of belonging* (fuzziness, Zimmerman 1985, Feoli and Zuccarello 1991), *diversity* (entropy, Margalef 1958,1989, Rényi 1961, Pielou 1975, Orlóci 1991a), *mutuality* (information, Rényi 1961, Feoli et al. 1984, Orlóci 1991a, 2001b), and other similar things that the reader may wish to identify.

7. *Probability scales: theoretical or empirical?* While theoretical probabilities come from basic principles, empirical probabilities are based on observation and experiment. My first preference is for empirical probabilities because these maximise the local relevance of the probability scale.

Many interesting points are made about randomisation experiments in the literature. The reader is referred to Edgington (1987) for theory and application as they relate to the traditional statistical models. For examples of ecological applications, I refer to Pillar and Orlóci (1996), McArdle and Anderson (2001), and to references in works quoted by these authors. One often reads text that makes much ado about the topic, and one discovers the confusion about a seemingly unimportant, but in actual fact highly consequential dichotomy. This is in the manner of incorporating randomisation testing into data analysis. The cases of the standard multivariate analysis of variance (MANOVA), and a similar analysis but formulated in cluster analytical terms provide a good basis for discussion:

1) Scientist SA elects to follow tradition and sets up the analysis in the usual terms of a MANOVA. This user ends up applying the test criterion $\theta_s = c_s/[1-c_s]$, which Morrison (1976, p. 178) explains. SA follows the usual steps of MANOVA, except the last when SA resorts to randomisation experiment to determine the Type I error probability of criterion θ_s.

2) Scientist SB proceeds in a different way, opting for the cluster analytical analogue of MANOVA as used by Edwards and Cavalli-Sforza (1965). SB's choice for a test criterion is the F ratio. This scientist uses randomisation testing to determine the Type I error probability associated with criterion F. SB's approach is not equivalent to SA's, but is it better or worse?

3) Suppose that the probability obtained by SA is smaller than the probability obtained by SB. An excited SA hastily concludes "the method of SB inflates the Type I error probability". But SA's exuberance could be short-lived. It is at the mercy of how the following question is answered: Does criterion $\theta_s = c_s/[1-c_s]$ describe reality? What are the points mitigating in the negative? The reader should note that the choice of θ_s confines the problem to a single characteristic axis of the $\mathbf{HE}^{-1}$ matrix (see Morrison 1976) whose largest characteristic root is θ_s. In all likelihood, θ_s will leave much of the residual variation unaccounted for in all but the most trivial cases of ecological practice. SA's problem could be confounded even further by a likely non-linear phase space configuration (see above) that cannot be handled effectively in terms of a characteristic root or characteristic axis of $\mathbf{HE}^{-1}$.

8. *Taxa: species-based or species-free*? Floristic heterogeneity in the sampling sites undermines the utility of species-based taxa for comparative purposes. The problem of comparisons is mitigated by use of taxa whose definition is consistent with the functionality of the individuals. *Cercidium microphyllum* (Leguminosae, Sonoran Desert of the U.S.A. and Mexico) and *Bulnesia retamo* (Zygophyllaceae, the Monte of Argentina) are typical examples. These species are distant in inheritance and geography, yet very similar in functionality (Orians and Solbrig 1977). A descriptive system that combines functionally identical species into the same taxa has high powers for description of reality and is particularly well suited for the purposes of comparative analysis in environment-related studies. Such a system is character based, and the characters may be sequentially or hierarchically arranged (see Podani 1985, Orlóci and Orlóci 1985, Orlóci 1991b, Pillar and Orlóci 1993, and references therein). The relevé from the Big Island of Hawaii, given in Table 2, is an example of hierarchical arrangement. The hierarchical nature of the relevé is owing to character nesting, implicit in character order. The nested arrangement allows the limits of "natural" taxa to be narrowed or broadened, and importantly, it allows also comparative analyses to be performed on any hierarchical level, any combination of levels, or any levels pooled. A comprehensive exposition of the schemes of character-based community analysis is found in a monograph by Pillar and Orlóci (1993).

Table 2. Partial record of Márta Mihály's Relevé #24 from the Devastation Trail area on the side of a cinder cone in Volcanos National Park, Hawaii. The character set is a selection from the original list. Date: 2000.01.15. Elevation: 1200 m. Land type: 40-year old cinder slope. Exposure: S. Slope: 15^0. Bare ground: 60% . Vegetation type: *Metrosideros-Cybotium-Styphelia*.

Characters \ Species	*Styphelia*	*Vaccinium*	*Dubautia*	*Nephrolepis*	*Cibotium*	*Metrosideros*	*Andropogon*	*Psilotum*	*Myrica*
Life-form (6 states)	2	2	3	3	1	1	4	3	2
Stem* tissue (4)	3	3	3	3	1	3	4	3	3
Leaf texture (4)	3	3	3	3	3	3	2	4	3
…									
Cover %	5	< 1	10	< 1	5	20	45	< 1	< 1

* or stem like organ.

9. *Prediction: the method of analogues*. The reader is referred to examples of prediction by analogues described in Box (1981), Küchler (1990), and Aszalós and Horváth (1998).[9] The method assumes that survival of specific plant forms in a site is not a straight chance event. Therefore, prediction is inherently a probabilistic exercise in which a choice is made among possible outcomes. But the outcome chosen for being most probable also has a certain probability of not being the one that actually materialises.[10] The *modus operandi* of predicting, say, diversity based on the method of analogues may go like this:

1) Determine the probabilistic linkage $p_{v|e}$ of each community type *v* (phytosociological association) to a specified type of the environment *e*.

2) Community type *v* that has highest probability $p_{v|e}$ to environmental type *e*, that has the highest probability p_e of occurring at a future point of time on a specific site, or at a different site at the same time, is the predictor community. The diversity of the predictor community is a prediction. The probability of the prediction being right is $p_{v|e}\, p_e$.

Epilogue

The daring, he who chooses to venture into the ecologist's world from an other field, soon becomes aware of layers of notions, facts, and believes from different historic times coexisting in the present. In other words, the present practice of ecology is a cross section of the past.[11]

But in the conceptual debris there is to be found constant things passed from one era to the next: interest in complex, dynamic things, in the pursuit of multidisciplinary approaches, and in dialectic reasoning about reality. To be noticed is the total absence of a purely mechanistic, recipe-driven ecological approach. What does this mean for Statistics? Simply, a statistical dialect to have utility in ecology must fall in with the ecological tradition. This makes sense, since Statistics is to serve ecology and not the other way around.

9   Other predictive models are used as well in ecology. Particularly to be noted is the Markov chain for which the reader finds examples and sources in Orlóci (2000, 2001b).

10   Chance interplaying with the deterministic laws of adaptation and ecological plasticity may be in the manner of randomness infesting determinism. This is a case of stochastics and it have to be treated as such (Orlóci 2001a). It is not unconceivable that determinism could arise from complete randomness, the case of the fractal (Mandelbrot 1977), albeit not a likely mechanism on which to base the prediction.

11   To see the cumulative process unfolding through time, the student may start with Kerner's literary piece "*Das Pflanzenleben der Donauländer*" (1863), continue with the literate classics like Clements (1916), Braun-Blanquet (1927), Gleason (1926), Post (1946), Whittaker (1962,1967), and Odum (1969), then progress on to readings about the origins of present interest lines in Lotka (1925, 1934), Voltera (1926), Watt (1947), Brey and Curtis (1952), Williams and Lambert (1959), Margalef (1958, 1968, 1989), Greig-Smith (1957, 1980), McIntosh (1967, 1985), Pielou (1969), Goodall (1953, 1970, 1972), Hill (1973), Maynard Smith (1974), Orlóci (1978, 1991a,b, 2000), May (1981, 1987), Shugart (1984), Svirzhev (1994), Feoli and Orlóci (1991), Pillar and Orlóci (1993), Solomon and Shugart (1993), Schneider (1994), and Czárán (1998), working his way back to more significant works through the references.

comments, and an anonymous reviewer for the same. I offer thanks in a very special way to Márta Mihály for her advise and participation in the completion of the manuscript.

References

Aszalós, R. and F. Horváth. 1998. Prediction of vegetation pattern on the regional scale. (In Hungarian) In: G. Fekete (ed.), *The Front Lines of Community Ecology.* Scientia, Budapest. pp. 161-170.

Berry, Th. 1990. *The Dream of the Earth.* Nature and Natural Philosophy Library, Sierra Club Books, San Francisco.

Box, E.O. 1981. *Macroclimate and Plant Forms: An Introduction to Predictive Modelling in Plantgeography.* W. Junk bv, The Hague.

Braun, E.L. 1950. *Deciduous Forests of Eastern North America.* Blackstone Co., Toronto.

Braun-Blanquet, J. 1927. *Pflanzensoziologie. —1932. Plant Sociology.* (Translated by G.D. Fuller and H.S. Conard.) McGraw-Hill, New York.

Bray, J.R. and J.T. Curtis. 1952. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* 27: 325-349.

Cain, S.A. 1944. *Foundations of Plant Geography.* Harper, New York.

Clements, F.E. 1916. *Plant Succession: An Analysis of the Development of Vegetation.* Publ. Carneg. Instn. 242.

Czárán, T. 1998. *Spatiotemporal Models of Population and Community Dynamics.* Chapman and Hall, London.

Delcourt, P.A. and H.R. Delcourt. 1987. *Long-term Forest Dynamics of the Temperate Zone.* Ecological Studies 63, Springer, New York.

Edgington, E.S. 1987. *Randomization Tests.* 2nd ed. Marcel Dekker, New York.

Edwards, A.W. F. & L.L. Cavalli-Sforza. 1965. A method for cluster analysis. *Biometrics* 2: 362-375.

Feoli, E. and L. Orlóci. 1991. (eds.), *Computer Assisted Vegetation Analysis.* Kluwer, Dordrecht.

Feoli, E. and V. Zuccarello. 1991. Syntaxonomy: a source for useful fuzzy sets for environmental analysis? In: E. Feoli and L. Orlóci (eds.), *Computer Assisted Vegetation Analysis*, Kluwer, Dordrecht. pp. 265-271.

Feoli, E., M. Lagonegro, and L. Orlóci. 1984. *Information Analysis of Vegetation Data.* W. Junk bv, The Hague.

Gleason H.A. 1926. The individualistic concept of the plant association. *Bull. Torrey Bot. Club* 53: 7-26.

Goodall, D.W. 1953. Objective methods for classification of vegetation. I. The use of interspecific correlation. *Aust. J. Bot.* 1: 39-63.

Goodall, D.W. 1970. Statistical plant ecology. *Ann. Rev. Ecol. Syst.* 1: 99-124.

Goodall, D.W. 1972. Building and testing ecosystem models. In: N.R. Jeffers (ed.), *Mathematical Models in Ecology*, Blackwell, Oxford. pp. 173-214.

Greig-Smith, P. 1957. *Quantitative Plant Ecology.* Butterworths, London.

Greig-Smith, P. 1980. The development of numerical classification and ordination. *Vegetatio* 42: 1-9.

Hill, O.H. 1973. Diversity and evenness: a unifying notion and its consequences. *Ecology* 54: 427-432.

Horváth, F., Z. Korsós, E. Kovácsné Láng, and I. Matskási (eds.) 1997. *Handbooks of the Hungarian National Biodiversity Monitoring System.* (In Hungarian). Hungarian Natural History Museum, Budapest.

Juhász-Nagy, P. 1994. *The Disappearing Diversity.* (In Hungarian) Scientia, Budapest.

Kerner von Marilaun, A. 1863. *Das Pflanzenleben der Donauländer.* Innsbruck, Wagner.

Küchler, A.W. 1990. Natural vegetation. In: E.B. Espenshade and J.L. Morrison (eds.), *Goode's World Atlas,* 18th ed. Rand McNally, Chicago. pp. 16-17.

Lean, G. & D. Hinrichsen. 1992. *Atlas of the Environment.* Helicon, Oxford.

Levin, S.A. (ed.) 2001. *Encyclopaedia of Biodiversity.* Princeton University Academic Press. San Diego.

Lotka, A.J. 1925. *Elements of Physical Biology.* Williams and Wilkins, Baltimore.

Lotka, A.J. 1934. *Théorie analytique des associations biologiques.* Hermann, Paris.

Manabe, S., K. Bryan and M. J. Spelman. 1990. Transient response of a global ocean-atmosphere model to a doubling of atmospheric carbon dioxide. *J. Phys. Oceanogr.* 20:722-749.

Mandelbrot, B.B. 1977. *Fractals: Form, Chance and Dimension.* Freeman, San Francisco.

Margalef, R. 1958. Information Theory in Ecology. *General Systems* 3: 36-71.

Margalef, R. 1968. *Perspectives in Ecological Theory.* University of Chicago Press, Chicago.

Margalef, R. 1989. On diversity and connectivity as historical expressions of ecosystems. *Coenoses* 4: 121-126.

Mason, J. 1990. The greenhouse effect and global warming. Information Office, British Coal, C.R.E. Stoke Orchard, Cheltenham, Gloucestershire, U.K. GL52 4RZ.

May, R.M. (ed.) 1981. *Theoretical Ecology.* Blackwell, Oxford.

May, R.M. 1987. Chaos and the dynamics of biological populations. *Proc. Royal Soc. London, Series A* 413: 27-44.

Maynard Smith, J. 1974. *Models in Ecology.* Cambridge University Press, London.

McArdle, B.H. and M.J. Anderson. 2001. Fitting multivariate models to community data: a comment on distance based community analysis. *Ecology* 82: 290-297.

McIntosh, R.P. 1967. The continuum concept of vegetation. *Bot. Rev.* 33: 130-187.

McIntosh, R.P. 1985. *The Background of Ecology: Concept and Theory.* Cambridge University Press, New York.

Milankovitch, M.M. 1941. Canon of isolation and the Ice-Age problem. *Royal Serb Acad. Spec. Publ.* 133.

Morrison, D.F. 1976. *Multivariate Statistical Methods.* McGraw-Hill, New York.

Odum, E.P. 1969. The strategy of ecosystem development. *Science* 164: 262-270.

Orians, G.H. and O.T. Solbrig. 1977. *Convergent Evolution in Warm Deserts.* Dowden, Hutchinson & Ross, Stroudsburg, Pennsylvania.

Orlóci, L. 1978. *Multivariate Analysis in Vegetation Research.* W. Junk bv, The Hague.

Orlóci, L. 1980. Non-linear data structures and their description. In: L. Orlóci, L. R. Rao, and W. M. Stiteler (eds.), *Multivariate Methods in Ecological Work*, International Co-operative Publishing House, Burtonsville, MD. pp. 191-202.

Orlóci, L. 1991a. *Entropy and Information.* Ecological Computations Series (ECS) Vol. 3. SPB Academic Publishing bv, the Hague.

Orlóci, L. 1991b. On character-based community analysis: choice, arrangement, comparison. *Coenoses* 6: 103-107.

Orlóci, L. 1993. The complexities and scenarios of ecosystem analysis. In: G.P. Patil and C.R. Rao (eds.), *Multivariate Environmental Statistics*, Elsevier Scientific, New York. pp. 423-432.

Orlóci, L. 1994. Global warming: the process and its anticipated phytoclimatic effects in temperate and cold zone. *Coenoses* 9: 69-74.

Orlóci, L. 2000. From Order to Causes. A personal view, concerning the principles of syndynamics. Published at the web address: hyperlink "http://publish.uwo.ca/~lorloci/" http://sites.netscape.net/lorloci

Orlóci, L. 2001a. Pattern dynamics: an essay concerning principles, techniques, and applications. *Community Ecol.* 2: 1-15.

Orlóci, L. 2001b. Biodiversity and its measurement. (Manuscript.)

Orlóci, L. and M. Orlóci. 1985. Comparison of communities without the use of species: model and example. *Ann. Bot.* (Roma) 43:275-285.

Orlóci, L. and V. De Patta Pillar. 1989. On sample size optimality in ecosystem survey. *Biometrie-Praximetrie* 29: 173-184.

Pielou, E.C. 1969. *An Introduction to Mathematical Ecology*. Wiley, New York.

Pielou, E.C. 1975. *Ecological Diversity*. Wiley, New York.

Pillar, De Patta V. and L. Orlóci. 1993. *Character-based Vegetation Analysis: the Theory and an Application Program*. Ecological Computations Series (ECS): Vol. 5. SPB Academic Publishing bv, The Hague, The Netherlands.

Pillar, V. De Patta and L. Orlóci. 1996. On randomisation testing in vegetation science: multifactor comparisons of relevé groups. *J. Veg. Sci.* 7: 585-592.

Podani, J. 1985. Syntoxonomic congruence in a small scale vegetation survey. *Abstracta Botanica* 9: 99-128.

Podani, J. 2000. *Introduction to the Exploration of Multivariate Biological Data*. Backhuys Publishers, Leiden.

Poore, M.E.D. 1962. The method of successive approximation in descriptive ecology. In: *Advances in Ecological Research*. Vol. 1, Academic Press, New York. pp. 35-68.

Post, L. Von. 1946. The prospect for pollen analysis in the study of the earth climatic history. *New Phytol.* 45: 193-217.

Rényi, A. 1961. On measures of entropy and information. In: J. Neyman (ed.), *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley. pp. 547-561.

Sampford, M.R. 1962. *An Introduction to Sampling Theory with Applications to Agriculture*. Oliver and Boyd, Edinburgh.

Scheuring, I. 1993. Multifractality: a new concept in vegetation science. *Abstracta Botanica* 17: 71-77.

Schneider, D.C. 1994. *Quantitative Ecology: Spatial and Temporal Scaling*. Academic Press, San Diego.

Shugart, H.H. Jr. 1984. *A Theory of Forest Dynamics: the Ecological Implications of Forest Succession Models*. Springer, New York.

Solomon, A.M. and H.H. Shugart. (eds.) 1993. *Vegetation Dynamics and Global Change*. Chapman & Hall, New York.

Svirzhev, Yu. M. 1994. Nonlinear problems in mathematical ecology. Working Paper WP-94-71, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria.

Volterra, V. 1926. Variation and fluctuation of the number of individual species living together. *J. Cons. Perm. Int. Ent. Mer.* 3: 3-51. Reprinted in: Chapman, R.N. 19 3. *Animal Ecology*. McGraw-Hill, NY.

Walker, D.J. and N.C. Kenkel. 1998. Fractal analysis of spatio-temporal dynamics in boreal forest landscape. *Abstracta Botanica* 22: 13-28.

Watt, A.S. 1947. Pattern and process in the plant community. *J. Ecol.* 35: 1-22.

Whittaker, R.H. 1962. Classification of natural communities. *Bot. Rev.* 28: 1-239.

Whittaker, R.H. 1967. Gradient analysis of vegetation. *Bot. Rev.* 42: 207-264.

Wildi, O. and L. Orlóci. 1991. Flexible gradient analysis: a note on ideas and an application. In: E. Feoli and L. Orlóci (eds), *Computer Assisted Vegetation Analysis*, Kluwer, Dordrecht. pp. 249-254.

Williams, W.T. and J.M. Lambert. 1959. Multivariate methods in plant ecology. I. Association-analysis in plant communities. *J. Ecol.* 47: 83-101.

Zimmerman, H.J. 1985. *Fuzzy Set Theory – and Its Applications*. Kluwer-Nijhoff Publishing. Dordrecht.