Statistical Ecology, Environmental Statistics, and Risk Assessment

G. P. Patil
The Pennsylvania State University

1 INTRODUCTION

Statistical ecology and environmental statistics are in a take-off stage both for reasons of societal challenge and statistical opportunity. It is becoming clear that statistical ecology and environmental statistics are calling for more and more of nontraditional statistical approaches. This is partly because ecological and environmental studies involve space, time, and innovative sampling and monitoring. Also, statistical ecology and environmental statistics must satisfy public policy responsibility in addition to disciplinary and interdisciplinary research. In conjunction with the 50th Anniversary of the International Biometric Society, it is only appropriate to attempt a historic perspective of statistical ecology and environmental statistics within a forward-looking biometric context.

The year 1994 marked the 25th year of statistical ecology and related ecological statistics with reference to the First International Symposium on Statistical Ecology held at Yale in 1969 with G. P. Patil, E. C. Pielou, and W. E. Waters as three co-chairmen representing the fields of statistics, theoretical ecology, and applied ecology. Over the past 25 years, statistical ecology has had a major impact on the collection, analysis, and interpretation of data on various fields of application and their theory. While much progress has been made in the past, the future promises even more rapid developments as sophis-
It is no wonder (Patil, 1995) that the Statistical Ecology Section of the International Association for Ecology and the related Liaison Committee on Statistical Ecology of the International Association for Ecology, the International Statistical Institute, and the International Biometric Committee have been around since their inception in 1969. And now the Ecological Society of America has a Statistical Ecology Section, and the American Statistical Association has a Section on Statistics and the Environment. The International Biometric Society and the American Statistical Association have together initiated a new Journal of Agricultural, Biological, and Environmental Statistics. For five years now, we have had the International Environmetrics Society with its journal, Environmetrics. And 1994 saw the creation of the Committee on Environmental Statistics at the International Statistical Institute and the launching of a new cross-disciplinary journal, Environmental and Ecological Statistics, published by Chapman and Hall. Also, an international initiative entitled SPRUCE (Statistics in Public Resources, Utilities, and in Care of the Environment) is now operational (Barnett and Turkman, 1993).

This space-limited overview necessarily has to be short and subjective. In this review article, we share some of the highlights and experiences in statistical ecology, environmental statistics, and risk assessment, giving references at the end for further reading. For purposes of organization, the sections are titled: ecological sampling and statistical inference; ecological diversity and its measurement; ecological assessments and generalized linear models; environmental policy and risk assessment; environmental data and cost-effective acquisition; landscape ecology and multi-scale assessment; and environmental monitoring, remote sensing, and geographic information systems.


2 ECOLOGICAL SAMPLING AND STATISTICAL INFERENCE

2.1 Encounter Sampling

Surveys for monitoring changes and trends in our environment and its resources involve some unusual conceptual and methodological issues pertaining to the observer, the observed, and the observational process (Southwood, 1968; Green, 1979; Waters and Resh, 1979; Patil, 1984). Problems that are not typical of current statistical theory and practice arise.

Traditional statistical theory and practice have been occupied largely with statistics involving randomization and replication. But in ecological and environmental work, observations most often fall in the nonexperimental, nonreplicated, and nonrandom categories (Hennemuth and Patil, 1983; Patil, 1991). Additionally, the problems of model specification and data interpretation acquire special importance and great concern. In statistical ecology and environmental statistics, the theory of weighted distributions (Rao, 1965; Patil and Rao, 1977; Patil et al. 1988) provides a perceptive and unifying approach for the problems of model specification and data interpretation within the context of encounter sampling.

Weighted distributions take into account the observer-observed interface, i.e., the method of ascertainment (Fisher, 1934; Haldane, 1938), by adjusting the probabilities of actual occurrence of events to arrive at a specification of the probabilities of those events as observed and recorded. Appropriate statistical modeling approaches help accomplish unbiased inference in spite of the biased data and, at times, even provide a more informative and economic setup (Patil and Taillie, 1989). For pseudo-replication, see Hurlbert (1984) and Gibbons (1994).

2.2 Adaptive Sampling

Several ecological and environmental populations are spatially distributed in a clumped manner. They are not very efficiently sampled by conventional probability-based sampling designs. Adaptive sampling is therefore introduced (Thompson, 1990) as a multistage design in which only the initial sample is obtained using a conventional probability-based procedure. When the variable of interest for a sampling unit satisfies a given criterion, however, additional units in the neighborhood are selected in the next sampling stage. This procedure is repeated until no new units satisfy the criterion, or the conditions of a stopping rule are satisfied. For methods of unbiased estimation and related statistical inference, see Thompson and Seber (1995).
With the recent growth of geographic information systems (GIS), spatial data for landscapes are becoming universal. This information provides a powerful aid to adaptive sampling and needs to be exploited.

2.3 Distance Sampling

Ecology is the study of the distribution and abundance of plants and animals and their interactions with one another and with their environment. Distance sampling theory extends the finite population sampling approach for purposes of estimating the population size/density. It is an extension of plot sampling methods, where it is assumed that all objects within sample plots are counted. As Seber (1993) puts it: In essence, one proceeds down a randomly chosen path called a line transect and measures/estimates the perpendicular distances from the line to the animals actually detected. Alternatively, one can choose a point instead and measure the radial distances of the animals detected. The methods apply to clusters of animals. At the heart of the methodology is a “detectability” function which is estimated in some robust fashion from the distances to the animals actually seen.

For more information, see Gates (1979), Buckland et al. (1993), and the extensive bibliographies in the two publications.

2.4 Capture–Recapture Sampling

The subject area of capture–recapture sampling has a long history in ecology, and has received a good deal of attention in the statistical and ecological literature. Much information is available on the size and dynamics of a population from repeat observations on identifiable individuals. As Cormack (1994) formulates it: Consider a series of s lists or samples in each of which a number of individuals are observed, and the marks are such that, at the end of the study the complete set of lists in which each individual is present can be formed without error... If the population is unchanging over the period of the study and if individuals independently have the same probability of appearing in any list, different from different lists, but unaffected by which other lists they appear in, this is the classic Petersen (with s = 2 samples) or Schnabel (s > 2) “census” analyzed by Chapman (1952), Darroch (1958), and many others.

For contingency-table and loglinear model approaches, see Fienberg (1972) and Cormack (1994). For closed populations, see Otis et al. (1978). For closed and open populations, see Pollock et al. (1990) and Seber (1973, 1992).

3 ECOLOGICAL DIVERSITY AND ITS MEASUREMENT

3.1 Introduction

Conservation biology, landscape ecology, and ecosystem-oriented natural resources management lend considerable urgency to issues and approaches concerning biodiversity assessment. Most of the traditional approaches and statistical tools are plot-based with a goal of definitive characterization. Diversity, however, is relative to spatial scale, temporal scale, and taxocene spectrum. Patterns may be more informative than absolutes in this regard.

The issues are fundamental in that explaining the effects of environment on the distribution and abundance of species is the essence of much ecological work. The controversies arise from the intrinsic scientific importance of diversity theories, as well as from the broad economic and social ramifications of considering biodiversity in land use decisions. At the heart of the scientific and social controversies regarding diversity are problems of quantification, interpretation, and analysis.

An important reason why current inference methods need improvement is that sampling for diversity is nonstandard. The more abundant a species is, the larger its range; or the more attracted it is to a light trap, the more likely it will be included in a sample. The smaller the extent or abundance of a species, the more likely it will be missing. The nonstandard sampling prevalent in diversity studies implies that off-the-shelf statistical techniques are frequently inappropriate.

3.2 Comparative Paradigms

Formulation and comparison of diversity measures have extensive precedents (Hurlbert, 1971; Kempton and Taylor, 1976; Pielou, 1975, 1977; Grassle et al., 1979; Magurran, 1988). This work, however, has mostly cast several measures as competitors rather than complementaries. The case basis for conventional diversity work lies primarily in local intensive studies, with recorded occurrence of taxa being considered definite and relative abundance estimates considered as quasi-ratio information. Issues of uncertainty, such as misidentification and differential detection, have been largely relegated to the background. Increasing representation of taxa with an expanding area of observation has been extensively studied, but issues of appropriate plot size and configuration have overshadowed the more fundamental implication that any diversity determination is relative to its areal basis. Temporal gap dynamics of forested landscapes are well-documented, but not explicitly recognized as a scale consideration in diversity assessments (Myers et al., 1995).

The classical view of diversity remains important for intensive studies of particular ecological communities and forest stands (Hunter, 1990; Swindel et al., 1991; Gove et al., 1994). However, the emerging sciences of landscape ecology and conservation biology have made evident the logistical and eco-
nomical impracticality of such intensive observational coverage for regions in the order of square kilometers and larger (Scott et al., 1989). These spatial scales are necessarily encompassed by contemporary ecosystem-oriented resource management and design of regional/national networks of biodiversity reserves. Furthermore, species/area and minimum viable population issues become fundamental in these matters.

3.3 Diversity Profiles

The multidimensional character of diversity can be revealed by establishing an intrinsic, and index-free, diversity ordering. A given pair of communities may not be comparable according to this intrinsic diversity ordering. In effect, diversity may appear to have decreased when viewed from one vantage point (i.e., index), and increased when viewed from a different perspective.

In view of the inadequacy of a single index, Patil and Taillie (1979, 1982) quantify diversity by means of diversity profiles. A diversity profile is a curve depicting the simultaneous values of a large collection of diversity indices. Thus, the profile portrays the views of diversity from many different vantage points simultaneously and in a single picture.

Differences in community diversity are studied by comparing profiles. If the two communities are intrinsically comparable, then one profile will lie uniformly above the other. Conversely, when the communities are not intrinsically comparable, their profiles may intersect. But even here, the profiles can reveal which portions of the community have undergone opposing diversity changes.

An important limitation of the traditional diversity analysis is its dependence upon relative instead of absolute abundance. Thus, processes which change total abundance without markedly affecting the pattern of relative abundance will not be detectable by a traditional diversity analysis. For broader community analysis, see Digby and Kempton (1987), Jeffers (1972, 1982), Legendre and Legendre (1987), and Reyment and Joreskog (1993).

4 ECOLOGICAL ASSESSMENTS AND GENERALIZED LINEAR MODELS

4.1 Introduction

Ecological scientists and data analysts have begun to find the traditional techniques of linear regression and analysis of variance somewhat limiting due to the requirements of constant variance and distributional assumption. Taylor’s power law (Bliss, 1941; Taylor, 1961) states that the variance increases with the mean. This is a robust generalization in ecology even if under vigorous discussion on its fine-tuning detail. Generalized linear models (GLM) extend the range of application by relaxing the requirements of classical linear models.

12 Statistical Ecology: Environmental Statistics, and Risk Assessment 219

It is encouraging to see that ecology seems to have begun to catch up with GLM. See, for example, Crawley (1993) and Kaur et al. (1995a,b).

Traditionally, transformation of the response variable has been the principal tool to remedy any evident breakdown. Witness the space devoted to data transformation in older biometry books, such as Bliss (1970). However, transformation of the response variable has its own limitations. Generalized linear models address these limitations, and the GLM framework thus allows the investigator to specify particular variance to mean relationships and draw inferences accordingly.

4.2 Generalized Linear Models and Ecological Applications

Logistic Regression

The binomial variance to mean relation leads to the well known logistic regression available for modeling binary data. Ecological applications include habitat association and resource selection (Manly et al., 1993; Ramsey et al., 1994), spatial association learning in hummingbirds (Graham and Petkau, 1994), bioassays with insects and insecticides for density dependence in mortality (Crawley, 1993), bioaccumulation (Futter, 1994), ecosystem modeling (Urfer, 1993), and others.

Probit Models

Probit analysis uses the inverse of the standard normal distribution function as its link function. In toxicology, the concept of tolerance distribution provides justification for the probit model. Ecological applications include beetles (Bliss, 1935), Daphnia magna (Sebaugh et al., 1991), insecticides and cockroaches (Hemingway et al., 1993), and forest pest management (Kreutzweiser et al., 1992).

Log-linear Models

The Poisson variance to mean relation leads to the log-linear model, a most effective approach for the analysis of count data (Bishop et al., 1975). Ecological applications include population size estimation by capture-recapture (Fienberg, 1972; Cormack, 1979, 1994), and geographical analysis of pollutants (Schwartz and Levin, 1991; Bailey et al., 1994).

Negative Binomial Variance to Mean Relation

The negative binomial distribution is frequently used to analyze count data when there are indications of overdispersion, possibly due to cluster sampling, spatial aggregation, etc. In fact, the variance to mean relation applies quite generally to Poisson mixtures. See Vogt et al. (1983) for an application to insect trap catches.
Constant Coefficient of Variation

For continuous nonnegative response data, the variance often increases with the mean and the distribution of the errors is generally skewed. Taylor’s power law suggests that the variance-mean relationship is approximately equivalent to a constant coefficient of variation. The gamma distribution with a fixed shape parameter is a linear exponential family having a constant coefficient of variation. The reciprocal link function is the canonical link for the gamma distribution, although the logarithmic link is often preferred because its range is the entire real line. Ecological applications relate to Drosophila melanogaster (McCullagh and Nelder, 1989), prey and predator (Crawley, 1993), and yield density experiments (Crawley, 1993).

Overdispersion Diagnostics

The presence of greater variation than expected for a nominal model is known as overdispersion. It is important to allow for overdispersion in the model in order to obtain correct variance estimates and valid hypothesis tests.

Many tests have been proposed for detecting overdispersion and for modelling any extra-variation detected in the data (Efron, 1986; Nelder and Pregibon, 1987). Ecological applications include fish toxicology (Ganio and Schaefer, 1992), rainfall and toxoplasmosis (Efron, 1986), and assessing toxicity of pollutants in aquatic systems (Bailer and Oris, 1994).

Generalized Estimating Equations

Liang and Zeger (1986) proposed this approach to deal with data in the form of correlated repeated measures (e.g., longitudinal data) in a semi-parametric framework. The basic idea lies in the generalization of the quasi likelihood estimating equations to allow a block-diagonal covariance matrix of the response vector. In a longitudinal data framework, this corresponds to the subjects being independent because of the design, but generating correlated observations. Ecological applications include spatial association learning in hummingbirds (Graham and Petkau, 1994), and teratology (Williams, 1975).

Survival Analysis

Sometimes ecological data have survival time as a response variable with censoring and with time-dependent covariates. See a study of parametric and semiparametric models for survival data involving Atlantic halibut in Smith et al. (1994) and a study involving lifetimes of seedlings in Crawley (1993).

5 ENVIRONMENTAL POLICY AND RISK ASSESSMENT

Environmental and ecological risk assessment has become an important area of study in response to environmental policy, planning, and evaluation (Cairns et al., 1976; Whyte and Burton, 1980; Bartell et al., 1992; Burgman et al., 1993; Suter et al., 1993; Jeffers, 1993). We wish to derive some feel for this vast topic using three contemporary investigations. For an excellent discussion of issues and approaches in environmental and ecological risk assessment, see U.S. National Academy (1993, 1994).

5.1 Crystal Cube for Coastal and Estuarine Degradation

Environmental decision makers would like to have a crystal ball predicting ecosystem response to stress. Instead, a crystal cube can be conceptualized, having a series of environmental indicator faces with each one in three colors representing no concern, warning, and alarm. In what follows is a statistical formulation based on O’Connor and Dewling (1986) in response to a question raised by the U.S. Ocean Dumping Act on what should constitute unreasonable coastal and estuarine degradation in light of field data: Let us call a measure of the pollutant effect on which an index is based an indicator variable. A measure of reproductive success in marine birds is an example. Broadly speaking, an indicator variable is usually a suitable nonnegative univariate summarization of a body of monitoring data. To be useful, it needs to have the following desirable properties:

1. Associatedness: It is associated with pollution, but is not necessarily a direct measurement of pollution itself.
2. Directionality: It increases with pollution. If a certain variable decreases with pollution, its reciprocal satisfies the directionality.
3. Sensitivity: It is sensitive to the causative pollution factor, so that the indicator variable has capacity to detect the pollution. Further, it is insensitive to the extraneous nonpollution, so that the indicator variable generates only limited “false alarms”.

Sometimes a question is raised as to the need for indices when we have good measures of pollutant effects. Why not just interpret the measurements when making decisions? The answer lies in the realization that useful interpretation of the measurement value does not usually come about until one has begun to see the measurement value as relative to some reference value.

We now wish to calibrate the indicator variable relative to some standard. But what yardstick do we choose and on what basis, so that the resultant indicator variable when measured in the units of the yardstick provides a useful index. O’Connor and Dewling (1986) propose that the yardstick be the benchmark or the reference value of the indicator variable that separates concern from no concern. The corresponding index then becomes

\[
\text{Index } (I) = \frac{\text{Indicator Variable}}{\text{Benchmark}}.
\]
Note that the benchmark is nothing but a critical value in some sense. Also note that regardless of the nature of the indicator variable and regardless of the actual choice of its benchmark, the index so defined has a most desirable common feature in that this definition of the index calibrates the index in terms of degradation with index value unity \((I = 1)\) separating concern from no concern.

How do we choose the errors of the first and second kind? We choose the error of the first kind to be 1 in 10, partly because ecological variability is rather large and also because a typical two-term ten-year manager may be able to stand a false alarm once in a ten-year period, which is also roughly half a human generation time. We choose error of the second kind to be 1 in 3 so that one is not caught napping in two successive years!

Finally, how do we define and choose a benchmark? See Patil (1991), Boswell et al. (1994), and their extensive references.

5.2 Causes of Fluctuation in Fishery Stock Sizes

The Chesapeake Bay Stock Assessment Committee was established in 1985 to access the fishery resources of the Chesapeake Bay. One of the initial tasks of the committee was to review and analyze available time series data related to fishery stocks and potential causes of trends in their abundance with the objective of relating the trends to changes in fishing and habitat variables.

Much of the variability in nonfishing mortality occurs during the prerecruit (Egg through Juvenile) stages. It is during these stages that one is most likely to be able to detect and establish pollutant and other environmental effects. While direct pollutant impacts upon adult stocks do not occur, such impacts tend to be nonlethal with long-term and subtle effects that are difficult to attribute to specific causes. For these reasons, modeling efforts have focused upon the early life stages.

Scatter plots of stock-recruitment data showed (Boswell et al., 1991) that some of the best year classes can occur when the stock is at a somewhat depressed level. This is often explained by asserting that the right combination of environmental factors occurred at precisely the right point in time and space. It follows that interactions among the explanatory variables are important. Indeed, the response may be “interaction-driven” to an extent that standard regression models may not adequately describe the nonlinearities. Modeling the interaction is further complicated by the need to consider where it occurs in space and time. Thus, low resolution measurements may miss or smooth out the interaction while high resolution measurements require too many parameters in the model.

Statistical problems that arise in connection with a conventional regression approach to the models include: (i) autocorrelated errors, (ii) lagged dependent variables, (iii) correlations between the errors and the explanatory variables, (iv) measurement errors, and (v) nonlinearities, including interactions and threshold effects, in the form of the response function.

12 Statistical Ecology, Environmental Statistics, and Risk Assessment

One approach that fitted well with the preceding biological models was that of categorical regression (Summers et al., 1984). The technique replaced all explanatory variables, including lagged dependent variables, by discretized binary versions. The Summers approach also assumes that any given combination of factors has the same effect at all lags.

The statistical literature has developed only partial answers to these problems. Nonlinearity may be handled by transformations, if appropriate forms can be found, or by nonlinear regression, if the functional form can be specified. If these approaches are not feasible, nonparametric methods or categorical regression may be considered.

The question of measurement errors has been examined in the context of structural and functional equation models. Typically, such methods are more limited in scope than linear regression methods and there is very little work on nonlinear models.

Models of lag structure have been developed in econometrics and in time series analysis. Transfer function analysis represents the most complete approach to the problem at this time. Again, linearity is assumed, although the measurement error problem has received some attention in the context of the Kalman filter.

By use of each of the above with appropriate theoretical extension where necessary, the relative importance of the issues raised earlier was assessed and a best compromise found for modeling stock variables. For details, see Boswell et al. (1991).

5.3 Implementation of the Benchmark Dose Method

The current approach to risk assessment for toxic noncarcinogenic chemicals is based on the assumption that there exists a threshold below which no adverse noncancer health effects are expected under lifetime exposure. Various regulatory agencies estimate a “safe” exposure by first determining an exposure level which has been shown to cause no adverse effect in animals or humans and then apply “uncertainty” factors to account for missing information.

Problems were identified with this methodology shortly after it was adopted some 30 years ago. The risk assessment community has been searching for improved methods since that time. One suggestion that has received a great deal of attention is to base the methodology on dose–response modeling. The idea is to estimate the effective dose (ED) that causes some critical effect in a specified percentage of the test animals (e.g., ED_{95} or ED_{90}) and then to designate the lower confidence limit for the effective dose as the “benchmark dose”. This benchmark dose may then be adjusted by uncertainty factors to arrive at the reference dose (RFD) or reference concentration (RfC).

In spite of the fact that it is generally agreed that the benchmark dose method addresses many of the shortcomings of the current methodology, more than a decade has passed since the benchmark dose method was suggested as an alternative. One reason for this delay is that there are a number of
difficult statistical issues remaining. While the potential benefits have been recognized, risk assessors have been understandably reluctant to adopt a methodology which is not yet completely understood.

There is a timely need to examine specific statistical and modeling issues that have been identified as impediments to widespread adoption of the benchmark dose method for risk assessment. The specific questions to be addressed are: (1) How should continuous response data be handled in modeling for the benchmark dose? (2) What is the best method for calculating a lower confidence limit in determining the benchmark dose? (3) Under what conditions might the estimated slope at the $ED_{10}$ (or $ED_{95}$) be useful in modifying the uncertainty factor?

For more details, see U.S. National Academy (1994).

6 ENVIRONMENTAL DATA AND COST-EFFECTIVE ACQUISITION

6.1 Observational Economy

Sampling consists of selection, acquisition, and quantification of a part of the population. While selection and acquisition apply to physical sampling units of the population, quantification pertains only to the variable of interest, which is a particular characteristic of the sampling units. Considerations of desirable criteria for representativeness and informativeness as variously defined usually lead to a desirable sample size of $n$ or more. On the other hand, considerations of resources in terms of cost, time, and effort usually lead to an affordable sample size of $n$ or less. A common experience is that $n << n$. This needs/resources dilemma has no universal panacea, but in appropriate circumstances, sampling protocols may be available that allow one to have both a large sample size and a small number of measurements, with all sampling units contributing to the information content of the measurements. We call this scenario “observational economy” (U.S. EPA 1995a, b). For observational economy to be feasible, a minimum requirement is that identification and acquisition of sampling units be inexpensive as compared with their quantification.

6.2 Design and Analysis with Composite Samples

Composite sampling has its roots in what is known as group testing. An early application of group testing was to estimate the prevalence of plant virus transmission by insects (Watson, 1936). In this application, insect vectors were allowed to feed upon host plants, thus allowing the disease transmission rate to be estimated from the number of plants that subsequently become diseased. Interestingly, the next important application of group testing seems to have occurred during World War II when U.S. servicemen were tested for syphilis by detecting the presence or absence of a specific antigen of the syphilis-causing bacterium in samples of their blood (Dorfman, 1943). In light of recent developments, composite sampling is increasingly becoming an acceptable practice for sampling soils, biota, and bulk materials. (see Gilbert, 1987).

If a composite measurement does not reveal a trait in question or is in compliance, then all individual samples forming that composite are classified as “negative”. When a composite tests positive, then retesting is performed on the individual samples or subsamples (aliquots) in order to locate the source of “contamination”. Generally, as the retesting protocol becomes more sophisticated, the expected number of analyses decreases. The analytical costs can be drastically reduced as the number of contaminated samples decreases.

A recent breakthrough with composite samples may be worth mentioning. The individual sample with the highest value, along with those individual samples forming an upper percentile, can now be identified with minimal retesting. This ability is extremely important when “hot spots” need to be identified such as with soil monitoring at a hazardous waste site (Gore and Patil, 1994). For more applications, see U.S. EPA (1995a).

6.3 Ranked Set Samples

Ranked set sampling is a little known method of sampling that allows the use of auxiliary information for improving upon the performance of simple random sampling. The primary requirement is the ability to rank sampling units with respect to the variable of interest without actually measuring that variable. Subjective judgment, prior experience, visual inspection, and concomitant variables are among the types of auxiliary information that may be used to achieve the ranking. The method does not prescribe any specific form or structure for the auxiliary information and the method is accordingly quite robust. Errors in ranking are permitted, although the better the ranking, the better the performance of the method.

Ranked set sampling (RSS), originally proposed by McIntyre (1952) and recently revisited by Patil et al. (1994), Kaur, et al. (1995b), and U.S. EPA (1995b), induces stratification of the whole population at the sample level, and provides a kind of double sampling estimator that is robust.

To see how RSS works, define a statistical sampling unit (ssu) to be a set of $m$ physical sampling units, where the sampling design parameter $m$ is the set size. A total of $n$ randomly chosen ssu are available for analysis, but only one physical unit is to be quantified from each ssu. The selection of this unit is the key to the ranked set sampling method. Let $r_1, r_2, \ldots r_m$ be positive integers with $r_1 + r_2 + \ldots + r_m = n$. All $n$ ssu are listed in a linear order at random. The lowest ranked unit is quantified in each of the first $r_1$ ssu. The second lowest ranked unit is quantified in each of the next $r_2$ ssu. This procedure continues until the highest ranked unit is quantified in each of the last $r_m$ ssu. In all, the ranked set sample consists of $n = r_1 + r_2 + \ldots + r_m$ quantifications of the available $nm$ units. The ranked set sampling design is said to be balanced if $r_1 = r_2 = \ldots = r_m = r$.

A real key to success may lie with ranking ability due to some “covariate” (David and Levine, 1972). For example, reflectance intensity of near-infrared
electromagnetic radiation, as recorded in a remotely sensed digital image, is directly proportional to vegetation concentration on the ground. Use of information from photographs and/or spatially referenced databases as found in a GIS can allow remote ranking prior to entering the field. (see Myers et al., 1994).

6.4 Sampling Heterogeneous Media

Regardless of the sampling design, when the medium is very heterogeneous, e.g., particulate material such as soil, ash, sludges, etc., the measurements on sample units may not be very representative. This can lead to bias, and also to high standard errors. The issue of representative sampling of particulate material has long plagued the mining industry (Taggart et al., 1945). In response, Pierre Gy (1982) has developed an applicable sampling theory. Regionalized variables by Matheron (1970) motivated Gy to consider two models, a continuous model which accounts for the continuous space or time variability of the characteristic of interest, and a discrete model which accounts for the discreteness of a population of fragments. He linked these two models with a short-range quality fluctuation term.

Basically, Gy identified all sources of error when obtaining a measurement of heterogeneous material, and derived the moments of these error distributions. His primary motivation was to define correct sampling and preparation of particulate material, so that all particles have an equal probability of being included in the final sample (see Bilionick, 1990; Pitard, 1993).

6.5 Combining Environmental Information

An increasingly important concern in environmental studies is the need to combine information from diverse sources that relate to a common endpoint and to combine environmental monitoring and assessment data as necessary and desirable. These are statistical problems, and statistical techniques are integral to analyses that combine such information/data.

Situations arise where a probability-based design (P-sample) is used to obtain unbiased estimates of population parameters such as the mean; however, some measurements may also be taken in a purposeful manner directed at suspected hot spots (non-P sample). The question is then, "May we combine the non-P sample data with the P-sample data? And how?"

The question has strong relevance for hazardous waste site monitoring, since there is often allowance for a certain amount of sampling to be taken at the field investigator's discretion where hot spots are suspected but were not chosen by a random P-sample. Also, with large-scale regional surveys of pollution, hotspot-directed sampling may have been done in the region over time for a number of reasons other than the regional survey. For more issues and approaches on combining environmental information, see Simberloff et al.

7 LANDSCAPE ECOLOGY AND MULTI-SCALE ASSESSMENT

Environmental and ecological data are available at a variety of spatial scales, arising from different sources which range from satellite imagery to field plots. Meanwhile, we need to make inferences about characteristics and processes at other scales, such as watershed boundaries, political subdivisions, or a desired unit of fixed size and shape.

The issue of measurement scale was addressed as Smith (1938) derived a regression relationship between the variance of crop yield and the size of quadrats. Greig-Smith (1964) analyzed contiguous quadrat data using nested analysis of variance. For an excellent discussion of the contemporary problem of pattern and scale in ecology, see Levin (1992a). For a statistical discussion, see Diggle (Chapter 18 this volume).

The effect of measurement scale on statistical inference in ecological studies is increasingly debated (Wiens, 1989; He et al., 1994; Schneider, 1994). According to Turner et al. (1989), a primary question is: What, if any, are the rules of extrapolating across scales?

Scaling through a fractal dimension can be used for estimating the spatial measure of an object. If the relationship between the log of estimated total length/area and the log of measurement scale is linearly decreasing, the estimate of the slope is used to estimate the fractal dimension. The fractal dimension increases with spatial complexity (Sugiwhara and May, 1990; Maurer, 1994). The spatial measure of the object is then predicted at scales other than the scale of actual measurement (Rastetter et al., 1992).

The relationship between measurements of scale-invariant systems manifests itself algebraically through a power law. For many natural phenomena, however, the exponent takes values that are not expected when measuring Euclidean objects. This was observed by Korcak (1938) when parameterizing the size distribution of islands in the Aegean Sea, and by Richardson (1961) when measuring the lengths of coastlines and other land frontiers. Mandelbrot and Wallis (1969) observed this with annual discharges of various rivers. It is in Mandelbrot (1983) that we see a recursive characterization in the form of a Koch Curve for natural boundaries, a Koch Snowflake for areas, and variations of Cantor Sets for size distributions. Mandelbrot calls them fractals and their dimension the fractal dimension, a Hausdorff dimension that strictly exceeds their topological dimension. Thus, fractal geometry seems to provide a more realistic characterization of the geometry of nature resultant from iterative and diffusive growth unconstrained by human manipulation.

For further ecological applications, see: Bradbury et al. (1984) for coral reefs; Morse et al. (1985) for arthropod habitat space on vegetation; Krummel et al. (1987) for landscape patterns distinguishing the natural from
the disturbed; Hastings and Sugihara (1993) for area-parameter relations; Costanza and Maxwell (1994) for landscape diversity; and Maurer (1994) for geographic range of breeding birds. Also see Simberloff et al. (1987) and Scott et al. (1993).

So far, results appear for additive variables where smaller-scale measurements/inference units are regular and hierarchically nested within larger scale units. Irregular, nonaligned multi-scale units present the greatest challenge; however, irregular hierarchically nested multi-scale units, such as the landscape units, seem to have promise of a tractable challenge. For a new interdisciplinary approach to understanding nonlinear ecological dynamics and its demographics, see Dennis et al. (1995).

8 ENVIRONMENTAL MONITORING, REMOTE SENSING, AND GEOGRAPHIC INFORMATION SYSTEMS

As environmental monitoring and management have grown increasingly ecosystem-oriented and landscape-sensitive, remote sensing and GIS have become core environmental information technologies (Goodechild et al., 1993; Fotheringham and Rogerson, 1994b; Michener et al., 1994; Glasbey and Berman, Chapter 19 this volume). Digital image data is the major domain of remote sensing. Digital image data is predominantly multivariate with sensor responses as variables and the pixels as observational cases. Major processing activities are image display, enhancement, and classification. Enhancement and classification are heavily statistical with local spatial autocorrelation present among pixels in a region.

Enhancement has both spatial and spectral concerns. Spatial enhancement necessarily involves interpolation. Interpolation involves nearest-neighbor, bilinear, or cubic convolution methods. Mostly ignored in these methodologies are issues of undersampling and oversampling that bear on the scale of spatial variation. Spectral enhancement variously involves kernel convolution filters, linear or nonlinear channel-composite “indices”, principal components, and other classical multivariate linear combinations. Except for the kernel-based convolution, spatial autocorrelation again tends to be ignored.

Two modes of automated theme generation are conventionally recognized. The “unsupervised” mode consists of cluster analysis to segment the image spatially, followed by sampling of each resulting cluster to determine its most appropriate thematic category label. In contrast, the “supervised” approach uses known instances of thematic categories in the image as “training sets” from which statistical “signatures” for categories are determined to drive a classification algorithm over the remainder of the image area.

There is a huge need for systematic, statistically based extraction of spatial structure from image data along the lines of local variogram analysis. This can only be forthcoming by close cooperation between statisticians and software engineers for remote sensing analysis systems.

Spectral distribution is usually viewed in terms of a (generalized) distance, with classifiers accordingly described as Euclidean distance, Mahalanobis distance, maximum likelihood, etc. Once again, local entropy properties of the image tend to be ignored in application of the classifier. Bayesian conditioning of classifiers has likewise been mostly global rather than local. Conditioning effects of known environmental subregions such as geology and soils have mostly been treated by various stratifying tactics rather than direct cognizance of the classifier. Until there is substantial work on direct incorporation of spatial autocorrelation and auxiliary conditioning in classifiers, the prospects of further progress are limited. Given recent developments in artificial intelligence, it appears that this work will involve the medium of neural networks.

Whatever form it takes, there is statistical opportunity for extending geostatistical ideas into the realm of classification. Unfortunately, it appears necessary to relax the usual geostatistical stationarity assumptions in such endeavors.

Geographic information systems are broadly defined as systems for input, storage, analysis, and output of geographically referenced information (see Maguire et al., 1991, for an extensive review). They are widely used as software toolboxes for integration of geographic data from multiple sources, including format change, resampling, interpolation, projection change, and restructuring; for implementation of many techniques of spatial data analysis (Fotheringham and Rogerson, 1994b); and for display of geographic data in map and related forms.

Because the sources of data integrated in a GIS application often vary in quality, there has been much work on the effects of data quality and uncertainty on the results of analysis (Heuvelink et al., 1989). Models of error for geographic data are complicated by the need to include spatial correlation terms (Goodechild et al., 1992; Openshaw et al., 1991).

Two schools of thought have emerged on statistical applications in GIS, broadly identifiable with the terms “spatial statistics” and “geostatistics”, respectively. The former, typified by the work of Anselin (1992), models spatial data sets as ensembles of interrelated but discrete objects, such as the patches in landscape models. The latter applies methods of geostatistics to GIS data sets conceived as digital representations of continuous fields. The future awaits substantial convergence between these two schools, and closer ties between GIS and statistical thinking in general. Extension of geostatistical ideas to multinomial fields, typified in GIS by data sets representing classifications of the land surface by soil type, land use, or vegetation class, will help this convergence, as will a clearer understanding of the circumstances in which it is better to model geographically distributed phenomena as fields, or as collections of discrete entities. For more information, see Sokal (1986), ter Braak (1986), Getis and Franklin (1987), Legendre (1993), Myers (1993, 1994), and Myers and Patil (1995).
9 LOOKING AHEAD

Statistical methods were initially developed for use in basic and applied sciences, and later in engineering and management. While basic statistical science is common to all areas, there are specific techniques developed to answer specific questions in each area. Ecological and environmental statistics is a relatively new subject and will need some of its own special methodologies. Typical ecological and environmental investigations are different from studies in the physical sciences and engineering. Unlike in the hard sciences, we have to deal with a longer span of investigations depending on life stages and their age lengths. Also, the instrumentation changes come by in response to the advancing technology.

Statistical thinking is an aid to the collection and interpretation of data. It may help clarify seeming confusion. It may help confuse seeming clarity. The statistical approach is expected to contribute to the overall balance, insight, and perspective of the substantive issue and its resolution in the light of the evidence on hand, be it in the nature of empirical data, literature-assembled data, expert opinion data, or a combination thereof.

Statisticians and biometricians have an important role to play in ecological research. And so also in environmental policy. As N. Phillip Ross (Ross, 1987), the Chief Statistician of the United States Environmental Protection Agency, has put it, "The statistician has to convince the decision makers that statistics can help them make better and defensible decisions and that statistical thinking is not so rigid as to preclude the flexibility in decision making." The subject area of ecological and environmental statistics offers this additional challenge and opportunity in the days ahead (see also Levin, 1992b).

ACKNOWLEDGMENTS

Dedicated to Professor C. R. Rao, my mentor, teacher, and colleague on his seventy-fifth birthday, September 10, 1995.

Prepared with partial support from United States Environmental Protection Agency under a Cooperative Agreement CR-821531. The contents have not been subjected to Agency review and therefore do not necessarily reflect the views of the Agency. The author is thankful to Michael Goodchild, Glen Johnson, Amarjot Kaur, Wayne Myers, William Stiteler, and Charles Taillie for helpful discussions in the preparation of this paper.

REFERENCES


12 Statistical Ecology, Environmental Statistics, and Risk Assessment


---

**12 Statistical Ecology, Environmental Statistics, and Risk Assessment**


