POTENTIAL BIASES IN SITE AND SPECIES SELECTION FOR ECOLOGICAL MONITORING

MICHAEL W. PALMER
Department of Botany, Oklahoma State University, Stillwater, OK 74078, U.S.A.

(Received: March 1992)

Abstract. A number of the contributions to this workshop (particularly in the session on biodiversity and population monitoring) have stressed the importance of choosing appropriate species and locations for ecological monitoring. This is of pragmatic necessity: it is impossible to monitor all species at all locations. The purpose of this note is to caution against some possible pitfalls in the selection of species and sites. These pitfalls are a result of the statistical concept of regression.

1. Introduction

Most scientists and resource managers are familiar with the use of regression for finding functional relationships between two or more variables. Here, I will use the term regression in the original sense, as coined by the 19th century statistician Sir Francis Galton. Upon examining a data set of more than 1000 fathers and sons, Galton noted that on average, very tall fathers had sons which were somewhat shorter, and very short fathers had sons who were somewhat taller. He considered this a regression to mediocrity, and defined the ‘Law of Regression’ by stating “Each peculiarity in a man is shared by his kinsman, but on the average in a less degree” (Quoted by Snedecor and Cochran, 1980).

The regression effect is found in many other kinds of data. For example, it is very common in examinations. According to Freedman et al. (1978), “In virtually all test–retest situations, the bottom group on the first test will on average show some improvement on the second test – and the top group will on average fall back.” We can think of each test score as consisting of ability plus luck (that is, a random element). The best scores on the first test in a large class will consist of very good students with very good luck. Since the element of “luck” is random, the initial top scores are unlikely (on average) to have quite as good scores on the second test. The opposite is true for the bottom scorers on the initial test.

I intend to demonstrate that the regression effect can crop up when monitoring rare species, if species selection is largely based on the species’ current rarity, or if site selection is based upon the abundance of the species being monitored.

2. Selection of Species

As discussed elsewhere in this symposium, a number of criteria are often used to select which species are most appropriate for monitoring. Frequently, one such criterion is the species’ current rarity. This is because we do not like to see rare
species becoming threatened, endangered, or extinct, and we would like to keep close watch on such potential problems.

One of the main objectives in monitoring is to be able to distinguish 'signal', or a directional trend, from 'noise', or random elements. Ecological data, including population numbers, are often exceptionally noisy: even if there is no net long-term trend, numbers can fluctuate dramatically. We do not wish falsely to conclude that a chance fluctuation is an actual trend.

I simulated random fluctuations of 20 species around each species' long-term average. Population data usually possess temporal autocorrelation: even in the absence of long-term trends, the number of individuals in one year is typically correlated with the number of individuals in the next. There are numerous ways to simulate chance fluctuations with temporal autocorrelation when no long-term trends actually exist; the detailed methodology on these are available from the author on request. The results of the simulations are given in Figure 1.

Despite the confusing jumble of Figure 1, it is evident that rare species tend to remain rare, common species tend to remain common, etc. However, chance variation assures that the species with the highest long-term average population size will not necessarily always be the most abundant species and, more importantly, the species which is rarest on the long term is not always the rarest at any given time. The net result is that if we use rarity at a particular instant of time as a criterion for monitoring, our species will on average appear to increase in abundance, even in the absence of real trends.