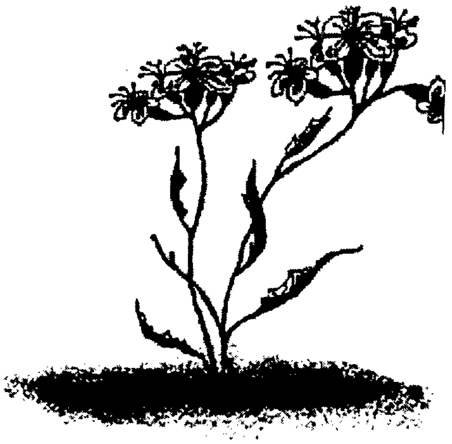
US Army Corps
of Engineers
Engineer Research and
Development Center

ERDC/CERL TR-00-5
February 2000

Stochastic Models of Plant Diversity: Application to White Sands Missile Range

Xiangchi Cao, George Z. Gertner, Bruce A. MacAllister, and Alan B. Anderson



Diversity is an average property of a community and is identified as the species variety and abundance. The study of biodiversity is important because it is one of the central themes in ecology; the diversity of a system is often seen as an indicator of the well-being of the system. In this study, we selected and defined theoretical statistical measures of plant diversity and developed theoretical dynamics models for plant communities. This project provides a new approach to measuring plant diversity through time.

The stochastic dynamics modeled in this project contain two main components: deterministic processes and stochastic processes.

Stochastic dynamics models make it possible to simulate plant diversity changes through time even without long-term observed data. Because both the time and space dimensions are included in these stochastic dynamics models, they have more extensive uses in assessing and monitoring plant communities. The potential applications of these models include (1) providing standard diversity measures, (2) monitoring the development of plant communities in terms of species diversity and structure diversity, (3) testing the significance of the influence of human activities on plant communities, and (4) estimating the rehabilitation time of a disturbed plant community.

WWW.CECER.ARMY.MIL/TECHREPORTS

Foreword

This study was conducted for the Office of the Directorate of Environmental Program (DAIM), Assistant Chief of Staff (Installation Management) (ACS(IM)) under Project 4A1622720A896, "Environmental Quality Technology;" Work Unit CN-T09, "Installation Capacity Factors." The technical monitor was Dr. Victor E. Diersing, DAIM-ED-N.

The work was performed by the Ecological Processes Branch (CN-N) of the Installations Division (CN), U.S. Army Construction Engineering Research Laboratory (CERL). The CERL Principal Investigator was Alan B. Anderson. Steve Hodapp is Chief, CECER-CN-N, and Dr. John Bandy is Chief, CECER-CN. The technical editor was Gloria Wienke, Information Technology Laboratory.

Part of this work was done by Xianchi Cao and Dr. George Gertner of the University of Illinois Department of Natural Resources and Environmental Sciences. Bruce MacAllister worked on the project as an Oakridge Associated Universities post-graduate research fellow. He also provided the cover graphic for the report.

The authors would like to thank White Sands Missile Range ITAM personnel for access to LCTA data and review of the report.

The Acting Director of CERL is Dr. Alan W. Moore.

CERL is an element of the U.S. Army Engineer Research and Development Center (ERDC), U.S. Army Corps of Engineers. The Acting Director of ERDC is Dr. Lewis E. Link and the Commander is COL Robin R. Cababa, EN.

Contents

List of Figures and Tables

Figures

Tables

1 Introduction

Background

The United States Army is responsible for managing over 12 million acres of land. The Army's land management objective is to maintain realistic military training and testing environments while promoting land stewardship. To accomplish this objective, the U.S. Army Land Condition Trend Analysis (LCTA) program was developed at the U.S. Army Construction Engineering Research Laboratory (CERL) under the sponsorship of the former U.S. Army Engineering and Housing Support Center (USAEHSC) as a means to inventory and monitor natural resources on military installations. LCTA uses standard methods to collect, analyze, and report natural resources data (Diersing, Shaw, and Tazik 1992) and is the Army's standard for land inventory and monitoring (Technical Note [TN] 420-74-3). Over 50 military installations and training areas in the United States and Germany have begun or plan to implement LCTA. LCTA data sets currently exist for more than 40 installations and contain 1 to 10 years of monitoring data. Lands inventoried as part of the LCTA program include Army Materiel Command (AMC), Forces Command (FORSCOM), Training and Doctrine Command (TRADOC), and National Guard Bureau installations. Over 75 percent of the Army's land base is represented by LCTA data (Shaw and Kowalski 1996).

An informal review of installation ITAM personnel indicated an interest in estimating plant diversity using LCTA data and modeling changes in plant diversity that result from alternative land uses.

Objectives

The objective of this project was to develop and test methodology to model changes in plant diversity using standard data from the U.S. Army's Land Condition Trend Analysis (LCTA) program. Specifically, stochastic models (those involving random variables) of plant diversity were to be developed using data from the White Sands Missile Range, New Mexico, LCTA program.

Approach

A literature survey was conducted to identify methods of characterizing and modeling plant diversity. Based on results of the literature review, modeling techniques were selected to model plant diversity. Land Condition Trend Analysis data from the White Sands Missile Range, New Mexico, was then used to develop stochastic plant diversity models for selected plant communities.

Scope

The modeling techniques described in this report are applicable to any military installation collecting LCTA data. The specific model documented is only applicable to White Sands Missile Range.

Mode of Technology Transfer

The information in this report will be provided to Army ITAM personnel responsible for using LCTA data. The information will also be provided to organizations responsible for developing and refining natural resources conservation methodologies through hard copy reports and through the CERL web site.

This study is part of a larger research effort to develop and field LCTA-related data applications. Models and analysis techniques described in this report are being incorporated into other modeling efforts and systems.

2 Ecological Diversity and its Measurement

Although the study of ecological diversity is only about 30 years old, it has been extensively represented in the literature during that time. There are three reasons ecologists are so interested in this topic. First, diversity is one of the central themes in ecology. Ecology is the scientific study of the distribution and abundance of organisms and the interrelationship between the organisms and their surroundings. Diversity studies, which address the variety and abundance of organisms, represent a major field of study in ecology. Second, diversity is often seen as an indicator of the health of an ecosystem. Studies have shown that pollution and disturbance reduce the richness and variety of the natural ecological communities. The loss of natural habitat and species extinction around the world have served to focus international attention on the issue of diversity. Third, considerable debate surrounds the measurement of diversity. On the surface, biodiversity seems to be a straightforward concept. Most people have an intuitive sense of the word. They would acknowledge that tropical rain forests harbor more species than temperate woodlands and are therefore more biologically diverse. However, the more we look at diversity, the less clearly defined it appears to be because diversity can be measured in so many different ways. A more in-depth study of diversity could reveal new and unexpected relationships between species and ultimately lead to a better understanding of the mechanisms involved. The study of ecological diversity over the past 30 years has raised three main questions: what is diversity, how is it measured, and what is its value in practice?

The Concept of Diversity and its Measurements

Diversity is one property of a biological community and consists of two components: variety and abundance. A large number of diversity measures have been devised by interpreting the relationship between variety and abundance in different ways. Magurran (1988) divides the measurement of species diversity into three main categories. First are the species richness indices. These indices are essentially measures of the number of species in an ecosystem. The second category of diversity measures includes those models that describe the distribution of species abundance. The third category is the diversity indices based on the relative species abundance. These indices, like the Shannon index and the Simpson index, consolidate species richness and evenness into a single figure.

Indices of Species Richness

Species richness is defined as the number of species or species density in the community. If a complete catalogue of species in a community can be obtained (it is possible for small communities), the number of species provides some measure of understanding diversity. Because most natural communities are very large, however, only a limited number of species can be counted and identified by sampling. Therefore, species density (defined as the number of species per unit area, biomass, or number of individuals), is commonly used as species richness (Hurlbert 1971; Homer 1976; Kempton and Wedderburn 1979; Kershaw and Loony 1985). Species density estimated by sampling varies with sample size and sample distribution. To cope with this problem Sander (1968) devised the technique of Rarefaction for calculating the number of species expected for all samples with standard sample size. Hurlbert (1971) improved Sanders' Rarefaction to produce an unbiased estimate of the number of species. Instead of using the number of species or the species density, others have used some simple indices derived from a combination of the number of species (S) and the number of total individuals (N) to represent species richness. Such indices include Margalef's index (Clifford and Stephenson 1975) and Menhinck's index (Whittaker 1970). The Margalef index (D_{Mg}) and Menhinck index (D_{Mn}) are, respectively, defined by:

$$D_{Mg} = \frac{S-1}{\ln N} \tag{1a}$$

$$D_{Mn} = \frac{S}{\sqrt{N}} \tag{1b}$$

Species Abundance Models

Species richness may be intuitive and easy to calculate but it does not contain any information of the relative abundance or distribution of species. In fact, species distributions are often more meaningful in explaining natural communities. Kempton and Wedderburn (1979) observe that a distribution of species is often a more sensitive measure of environmental disturbance than species richness alone. In an early study, Fisher, Corbet, and Williams (1943) found that patterns occur in species distribution. It is very rare to have an equal number of individuals for all species. Instead, a few species would be very abundant, some would have a medium abundance, while most species have only a few individuals. This observation led to the development of species abundance models.

Pielou (1975) developed six models of species abundance. They are the niche preemption model, broken stick model, overlapping niche model, truncated negative binomial distribution, log-series distribution, and log-normal distribution. In a monograph of stochastic abundance models, Engen (1978) lists nine mathematical distributions related to species abundance. These are the gamma distribution, first kind beta distribution, second kind beta distribution, log-normal distribution, Dirichlet distribution, negative binomial distribution, logarithmic series distribution, negative binomial beta distribution, and Poisson log-normal distribution. A number of other species abundance models also appear in literature (Zipf 1965; Mandelbrot 1977; Gray 1988). In practice, four main models (Magurran 1988) characterize species diversity: the log-normal distribution, geometric series, logarithmic series, and MacArthur's broken stick model.

The geometric series is based on the hypothesis that every species ranked from the most to the least abundant take the same proportion (k) of the remainder. The ranked abundance list is k, k(1-k), $k(1-k)^2$, ... $k(1-k)^{S-2}$, $(1-k)^{S-1}$. May (1975) gave the probability distribution of such ranked list, F(x), which is defined as the probability that a randomly chosen species has size less than x.

$$(2) \qquad \mathrm{F(x)} = \mathrm{C} - \frac{\ln(x)}{\ln(1-k)}$$

C is a constant. Geometric series pattern of species abundance is found primarily in species-poor environments or in the very early stages of a succession (Whittaker 1965, 1970, 1972).

Fisher, Corbet, and Williams (1943) derived a log-series model to describe the species abundance of Malayan Lepidoptera. This log-series model represented the first attempt to describe mathematically the relationship between the number of species and the number of individuals in those species. In the log-series, the expected frequency of a species with abundance x is given by:

$$(3) \qquad f_X = \frac{ab^x}{x}, \text{ for } x = 1, 2, \ldots$$

The variable b ($0 < b < 1$) is a constant that is dependent on the sample size, and a ($a > 0$) is a constant determined by the characteristics of the community. When the abundance of each species is plotted on a logarithmic scale in rank, the log-series approximates a straight line with a slope of -a (Taylor, Kempton, and Woiwod 1976). The log-series provides a statistical satisfactory description of samples from a wide range of communities (Williams 1964; Kempton and Taylor

1974; Gray 1978). It can be used for small, stressed, or pioneer communities (May 1975; Whittaker 1965), and also for "neutral" unstressed communities (Caswell 1976).

Preston (1948, 1962) observed that in many samples middle-ranked species were relatively numerous, and there were fewer rare species than the log-series distribution predicted. By the log transformation of number of individuals, Preston found that the number of species always distributed with a truncated normal distribution. In a log-normal distribution, the probability density function is:

$$\text{(4)} \qquad \mathrm{f(x)} = \frac{1}{x\sqrt{2m*\sigma}} \exp\left[\frac{-(\ln x - m)^2}{2\sigma^2}\right]$$

The expected size of species is calculated by:

$$\text{(5)} \qquad \mathrm{E(S)} = \exp\left[m + \frac{\sigma}{2}\right].$$

The log-normal distribution rises from the statistical properties of large numbers and as a consequence of the Central Limit Theorem (May 1975). Thus, it is considered to be the descriptor of large and mature natural communities (May 1975; Whittaker 1972; Gray 1978; Preston 1980; Magurran 1988).

MacArthur (1957) first proposed the broken stick model. In this model, the resource is likened to a stick broken randomly and simultaneously into S distinct segments. The lengths of the segments represent the "sizes" of the species. According to the model, the expected size of the i-th species, xi, is:

$$\text{(6)} \qquad \mathrm{E(x_i)} = \frac{1}{S}\sum_{k=i}^{S} \frac{1}{k}$$

The broken stick model reflects a much more equitable state of affairs than those suggested by the log-normal, log-series, and geometric series. It has good fits for the communities with a few species and relative high evenness between species (May 1974; Pielou 1975).

Diversity Indices

The third kind of diversity measures include those indices based on the proportional abundance of species. Although species abundance distributions provide the fullest description of diversity data, there are times when a single diversity index is required. When communities do not fit one model and it is desired to compare them by means of diversity, indices based on the proportional abundance of species provide a solution to this problem. The most commonly used indices are the Shannon index and Simpson index.

The Shannon index is defined as:

$$(7) \qquad H' = -\sum_{i=1}^{S} (p_i \, ln p_i)$$

The parameter p_i is the proportion of community members in the i-th class. Shannon originally proposed this diversity index as a measure of the information content of a code. The Shannon index assumes that individuals are randomly sampled from an "infinitely large" population and all species are represented in the sample.

Using the fact that the probability of drawing two successive individuals belonging to the same species in a random sampling is p^2_i, Simpson (1949) suggested a statistic, D, that has the form of:

$$(8) \qquad D = \sum_{i=1}^{S} P_i^2$$

The parameter p_i is the proportion of individuals in the i-th species and S is the total number of species in the community. This statistic measures a property that is opposite to the diversity. The diversity index corresponding to the statistic D, the Simpson index, is then given by:

$$(9) \qquad H = -\ln D$$

The Shannon index and Simpson index are two special cases of a more general class of functions (H) used in mathematical theory of information (Pielou 1975).

$$(10) \qquad H_\alpha = \frac{\ln(\sum_{i=1}^{S} p_i^\alpha)}{1-\alpha}$$

$$(11) \qquad H_1 = \lim_{\alpha \to 1}(H_\alpha) = -\sum_{i=1}^{S} p_i \ln(p_i) \qquad \text{(Shannon index)}$$

$$(12) \qquad H_2 = H_{\alpha=2} = -\ln(\sum_{i=1}^{S} p_i^2) \qquad \text{(Simpson index)}$$

Using the concept of rarity, Dennis and Patil (1979) find a class of diversity functions that also leads to the Shannon index and Simpson index.

(13) $$\text{Rarity} = R(p_i) = \beta^{-1}[1 - p_i^b]$$

(14) $$\text{Diversity} = \Delta_\beta = \sum_{i=1}^{S} p_i R(p_i)$$

(15) $$\Delta_{-1} = \Delta_{\beta=-1} = S - 1$$ (Species richness index)

(16) $$\Delta_0 = \lim_{\beta \to 0}(\Delta_\beta) = -\sum_{i=1}^{S}(p_i \ln p_i)$$ (Shannon index)

(17) $$\Delta_1 = \Delta_{\beta=1} = 1 - \sum_{i=1}^{S} p_i^2$$ (Simpson index)

There are other diversity indices such as the McIntosh index and Berger-Parker index. McIntosh (1967) found that the Euclidean distance of the assemblage from the origin could be used as a measure of diversity. Berger and Parker (1970) use the proportional importance of the most abundant spices as the diversity measure.

New Models of Diversity

Hughes (1984, 1986) has completed research on the diversity measures with a dynamics model. In this dynamics model, the abundance (n) of the i-th species at time t+1 is calculated from the expression:

(18) $$n_{t+1}^{i} = S[n_t^i + R(1 + Zn_t^i)\frac{K - N_t}{K}]$$

The variable S is the survival parameter, R is the recruit parameter, Z is the aggregation parameter, and N_t is the number of total individuals. This dynamics model simulates the development and the progression of a theoretical community through time. It defines a “community” with a variable number of species and species abundance.

Diversity itself is also divided into two types: species diversity and spatial diversity (structure and habitat diversity). The diversity discussed above is mainly concerned with species diversity. Although species diversity is the more important, structure and habitat diversity has special use in ecology. Habitat diversity has been used as an important component of wildlife conservation evaluation (Fuller and Langslow 1986; Pearsall, Durham, and Eagar 1986; Usher 1986). MacArthur and MacArthur (1961) found that the structural diversity of a temperate woodland in North America was a much better predictor of bird diversity than the plant species diversity. Southwood, Brown, and Reader (1979) reported finding a close relationship between plant spatial diversity and insect diversity in woodland succession. Elton and Liller (1954) divided habitat diversity into four levels. The first level is the major habitat system (e.g., terrestrial or aquatic). The major habitat system is then divided into formation type (e.g., woodland or open land). When categorizing the formation type, the presence of vertical layers (e.g., ground flora, shrub, high canopy) is recorded. These vertical layers comprise the third level of habitat diversity. A fourth layer of qualifiers (e.g., conifer, deciduous) then describes them further.

Different situations and studies may have different habitat classification schemes. Once the structure and habitat diversity is defined, the next question is how to measure the structure and habitat diversity Methods for measuring species diversity can also be used for measuring structure and habitat diversity (Magurran 1988). A rather different approach, differentiation diversity, is required when we wish to ascertain how species numbers and identifiers differ between communities or along gradients (Magurran 1988).

The true value of studying diversity is in its application. It is believed that diversity is a good indicator of the well-being of an ecosystem (Magurran 1988). Diversity measures have potential applications to two main ideas. First, is the idea of natural resources conservation, which is underpinned by the notion that species-rich communities are better than species-poor ones. Second is the idea of environmental monitoring, which assumes that the adverse effects of pollution or disturbance will be reflected in a reduction in diversity (or by a change in the shape of the species abundance distribution [Magurran 1988]). Many researchers (Bechtel and Copeland 1970; Schafer 1973; Rose 1978; Gray and Mirza 1979; Yapp 1979; Wu 1982; Usher 1986; Tomascik and Sander 1987) have shown successful applications of diversity measures.

It should be noted, however, that all of the previously mentioned diversity measures have some limitations. Species richness gives only the number of species or species density. Diversity indices based on the proportional abundance of species represent the number of species and their relative abundance as a single figure. However, they lose information about relative species abundance. Species abundance models give the fullest description of species distribution. Once the species distribution of a community is determined, diversity indices can be calculated from the distribution of species. However, species abundance models also have problems in application.

Although there are many mathematical distribution models related to species abundance, two models have focused on the limited distributions of species: the log-normal and the log-series models. The log-normal model is said to fit a wide variety of communities that are stable and contain a large number of species, while the log-series model is said to be applicable to small, stressed, or unstable communities. Although the log-normal and the log-series models have been widely used and have good theoretical explanations, there are three problems with using these models. First, the majority of real communities may not be satisfactorily described by either the log-normal or the log-series. It is very rare to find a community with the exact log-normal distribution or log-series distribution. Second, either the log-normal or the log-series distribution may result as an artifact of the sampling procedure. Combining smaller, separate samples may produce a log-normal distribution, even if the individual samples each show a log-series distribution. On the other hand, a small sample taken from a log-normal community may produce a distribution resembling a log-series. Third, even if the log-normal or the log-series distributions fit the community very well by the 'goodness of fit' test, the fitted log-normal or the log-series model does not provide a good estimation of the abundance of the most prevalent species or the least prevalent species. This is because the most abundant and least abundant species are represented at the tails of the distribution. Most biodiversity measures like the Shannon index and the Simpson index depend on a large number of rare species. Also, the quality of a community is mainly determined by the most abundant species. Therefore, using a theoretic distribution in biodiversity analysis causes a precision problem.

Although the problems discussed here are critical in a diversity study, they have not been solved yet. This study intends to explore a new method of measuring diversity and fill the gap in the recent diversity study. A new stochastic dynamics model will be developed to model both the deterministic population growth and population fluctuations. This model will use a variable number of species and species abundance to calculate any kind of diversity indices. Because both the dynamic changes and stochastic fluctuations are included in the new model, the stochastic dynamics model has the potential for extensive use in natural resources management and environmental monitoring. Future applications of these stochastic dynamics models include (1) providing standard diversity measures, (2) monitoring the development of plant communities in terms of species diversity and structure diversity, (3) testing the significance of the influence of human activities on plant communities, and (4) estimating rehabilitation time for a disturbed plant community.

3 Stochastic Dynamic Models

Although the real mechanisms for the development of a plant community are very complicated, mathematical models can replace the complex biological reality with some idealized hypothetical systems. In fact, many simplified mathematical models can interpret complex consequences and predict the behavior of elaborate natural systems. In this report, we will discuss several population dynamics models with environmental stochasticity and demographic stochasticity.

A stochastic dynamics model can be derived by manipulating a deterministic model to include demographic stochasticity and environmental stochasticity. A deterministic model describes the mechanisms controlling the population growth and decay. The values of the parameters in the deterministic model are assumed to be known. Whenever the initial conditions are given, the deterministic models provide exact predictions of future populations. In the real world, however, few communities have the population dynamics described by deterministic models. Instead, populations of most natural communities have fluctuating growth because of demographic and environmental stochasticity (or noise).

Turelli (1986) provides definitions of environmental stochasticity and demographic stochasticity. Demographic stochasticity (or within-individual variability) is the variation of individuals who appear to be identical but have different life lengths and produce different numbers of offspring. Integer-valued stochastic models are typically used to investigate the consequences of the demographic stochasticity. Environmental factors vary unpredictably through time in ways that affect all individuals. This variation is called environmental stochasticity. Most analyses of the consequences of environmental stochasticity begin by adding a noise term in the deterministic model (May 1973; Capocelli and Ricciardi 1974; Tuckwell 1974; Goel and Richter-Dyn 1974; Turelli 1977). This produces stochastic difference and differential equations with continuous ranges.

In this paper, we first introduce a dynamics model of population growth. This model derives birth and death rates as they relate to population growth from the relationships among plants. It also derives these rates from the relationship between plants and the environment. The model will serve as the deterministic part of the stochastic model. Next, we present some well-known stochastic models such as birth and death processes and a diffusion process. In this part, we demonstrate different methods to solve a variety of stochastic differential equations. Next, we will show how to incorporate both demographic noise and environmental noise into a single model that considers the joint effects of demographic and environmental stochasticity. Finally, we will apply the model to predict the dynamics of the understory in a bottomland forest.

A Simulation Model of Population Dynamics

The dynamics of the population growth of a community can be mathematically described if the functional behavior of the rate of growth is known. The literature contains many different representations of the growth rate (Nisbet and Gurney 1982; Streifer 1974; Hallam 1986). Most of these representations have two basic factors in common: survivorship and recruitment. Survivorship can be expressed by the percentage of individuals that survive from one time period to the next. Recruitment is the addition of new individuals to the population by immigrations and births. For plant communities, the emigration can be neglected because plants usually do not emigrate once they become established in the community.

A Discrete Model of Plant Population Dynamics

Consider a 1-hectare plant community consisting of 10,000 space units (i.e., 1 space unit = 1 m^2). Each individual of i-th species occupies A^i space unit. Thus, the space occupied by all plants at time t is $A_t = \Sigma x_t^i * A^i$, where x_t^i is the number of individuals of the i-th species at time t.

The change in the number of individuals of the i-th species from time t to time t+1 depends on the number present at time t, the number recruited (immigration and birth), and the number of individuals that survive. This relationship is expressed as:

$$(1)^* \qquad x_{t+1}^i = s^i (x_t^i + r^i)$$

where x_{t+1}^i and x_t^i are the number of individuals of the i-th species at time t+1 and t, respectively, s^i is the net survivorship, and r^i is the net recruitment.

---

* Equations in each chapter begin with (1).

Competition, insects and disease, senescence, and occasional catastrophe (natural catastrophe and human disturbance) are the most important sources of mortality for terrestrial plants. Mortality (or survivorship) is incorporated into the model in three ways. First, the inter-specific survive rate (S) is included in the model as a given number ranging from 0 to 1. The inter-specific survival rate of the i-th species (S^i) represents the intrinsic survivability of the i-th species. Second, to account for the mortality due to a catastrophe, the inter-specific survival rate is multiplied by the catastrophe index C ($0<C<1$). We assume that all species suffer from the catastrophe but their tolerances are different. Third, the risk of mortality from competitive interference is likely to increase with an increase in the community density. In this model we use a discrete-time analog of the Lotka-Volterra competition equation which is expressed as:

$$(2) \qquad Q^i = g_i\left(S\frac{a_{ij}x_t^j}{K^i}\right)$$

where K^i is the maximum number of individuals of the i-th species and a_{ij} are the competition coefficients. We assume that for each i, $g(x)>0$, and $g(x)'<0$ for $x>0$ (i.e., g(x) is a positive decreasing function). Thus, the survivorship component of equation (1) is now written as:

$$(3) \qquad s^i = \frac{S^i * C}{1+Q^i}$$

The number of individuals recruited depends on the inter-specific recruitment potential (R^i), the availability of space (A_t/A) at time t and the immigration (I). The inter-specific recruitment rate represents the rate of addition of population by birth. As the community develops and as space becomes occupied, the potential for further recruitment is proportionally reduced. Therefore, the inter-specific recruitment potential (R^i) should be multiplied by a space constraint (V) which is a function of the ratio of space occupied and total space, i.e., $V = f(A_t/A)$. The combined recruitment rate is:

$$(4) \qquad r^i = (R^i x_t^i + I^i)V$$

Putting s^i and r^i in equation (1), we obtain the final version of the dynamics model.

$$(5) \qquad x_{t+1}^i = \frac{S^i * C}{1+Q^i}\,[x_t^i + (R^i x_t^i + I^i)V]$$

This model produces a “community” with a variable number of species and variable population, which can then be used to calculate diversity indices.

The Continuous Model of Plant Population Dynamics

When the population size of a community is large, a continuous model can approximate population growth. Let x^i_{t+dt} and x^i_t respectively be the numbers of i-th species at times t+dt and t. The change in the number of individuals of the i-th species from time t to time t+dt is:

$$(6)\qquad x^i_{t+dt} - x^i_t = \{x^i_t[R^i \frac{S^i * C}{1+Q^i} V - (1-\frac{S^i * C}{1+Q^i})] + \frac{S^i * C}{1+Q^i} I^i V\}*dt$$

$$= [x^i_t(b^i-d^i)+m^i]*dt$$

Dividing dt on both sides of (6) and letting dt go to 0, we have the following differential equation.

$$(7)\qquad \frac{d}{dt}(x^i_t) = x^i_t(b^i-d^i)+ m^i$$

$$= x^i_t*h^i + m^i$$

where $b^i = (R^i \frac{S^i * C}{1+Q^i} V)$ is the birth rate, $d^i = (1-\frac{S^i * C}{1+Q^i})$ is the death rate, m^i is the net immigration rate and h^i is the net growth rate. The solution of equation (7) is an exponential growth function with immigration.

$$(8)\qquad x^i_t = (x^i_0 + \frac{m}{r})*\exp(h^i t) - \frac{m}{r}$$

When h>0, immigration causes an exponential increase in population growth. When h<0, immigration causes an exponential decrease in population growth, and there is high competition in the community. The new immigrants increase the competition and speed up the decrease of population growth. As in equation (5), equation (8) also produces a “community” with a variable number of species and variable population.

Demographic Stochasticity

As discussed in the dynamics models, the birth rate, death rate, and immigration rate are assumed to be constant. In the natural world, however, this assumption does not usually hold true. Individuals of identical plant types may have different life lengths and produce different numbers of offspring. This variation among the individuals is called demographic stochasticity. There is vast literature on the modeling of demographic noise. Turelli (1986) gives a very good overview of demographic stochastic models.

In this paper we will use the birth and death processes to describe the demographic stochasticity. The birth-death processes describe population dynamics with biologically accurate, interpretable birth and death rates and are applicable to individual numbers of every size. First, we start with an 'external birth' model, the Poisson process of immigration.

The Poisson Model of Immigration

Suppose the chance of an event (immigration) in the small time interval (t, t+dt) is λdt+o(dt), where the last term is a remainder term which becomes negligible compared with dt as dt gets smaller, and may be consequently neglected in comparison with the first term. This chance λdt is assumed to be independent of the number of previous happenings, and, moreover, each event is assumed to be independent. Therefore the chance of two events (or more) occurring in the time interval (t, t+dt) is o(dt), and is also negligible. Let the probability of r events (and no more nor less) in the time interval (0,t) be p_r. Then we can mathematically represent the whole set or distribution of probabilities p_r (r=0, 1, 2, ...) by the generating function in θ,

$$\Pi(\theta) = i = 0 \;\; p_r\theta^i \tag{9}$$

As the p_r are dependent on t, we note that Π(θ) is also a function of t, and shall write it more fully as $\Pi_t(\theta)$. Then by the rules of probability, as the increase in the total number of events, during the further interval dt, is independent of the previous total number occurring in the interval (0, t), it follows that

$$\Pi_{t+dt}(\theta) = \Pi_t(\theta)\{1-\lambda dt+\lambda\theta dt\} \text{ or}$$

$$\frac{\partial}{\partial t}\log(P_t(q) = \lambda(\theta-1) \tag{10}$$

whence, as $\Pi_0(\theta) = 1$,

$$\Pi_t(\theta) = \exp[\lambda t(\theta-1)], \text{ and} \tag{11}$$

$$p_r = (\lambda t)r{*}\exp(-\lambda t)/r! \tag{12}$$

Equations (11) and (12) specify the Poisson distribution with mean $m=\lambda t$. When λ is time-varying, the mean is obtained by the following equation:

$$m = \int_0^t \lambda(u)du \tag{13}$$

The Simple Birth and Death Process

We assume that the probability for any given individual to give birth in the time interval (t, t+dt) is ldt while the probability of dying in that time interval is mdt. Equivalently, denoting by dX(t, t+dt) the increment X(t+dt)-X(t) of the population size in (t, t+dt) we can make the following assumptions:

$$P\{dX(t, t+dt) = 1/X(t) = n\} = lndt + o(dt) \tag{14a}$$

$$P\{dX(t, t+dt) = -1/X(t) = n\} = mndt + o(dt) \tag{14b}$$

$$P\{|dX(t, t+dt)| > 1 / X(t) = n\} = o(dt) \tag{14c}$$

From (14) it follows

$$p\{dX(t, t+dt) = 0/X(t) = n\} = 1-(l+m)ndt + o(dt) \tag{15}$$

From (14) and (15) we easily obtain:

$$p_n(t+dt) = p_n(t)\,[1-(l+m)ndt] + p_{n-1}(t)(n-1)ldt + p_{n+1}(t)m(n+1)dt + o(dt)$$

or, the limit as dt approaches 0:

$$\frac{dp_n(t)}{dt} = -n(l+m)p_n(t) + l(n-l)p_{n-1}(t) + m(n+1)p_{n+1}(t) \tag{16}$$

with the initial condition:

$$p_n(0)=\begin{cases}1, n=j\\ 0, n\neq j\end{cases} \tag{17}$$

Before showing how equation (16) can be solved to determine the functions $p_n(t)$, let us calculate the mean population size $E[X(t)]$ and its variance $\mathrm{Var}\ \{X(t)\}$ by a straightforward procedure. Use the definition:

$$\text{(18)} \qquad E[X(t)] = n = 0 \ \ np_n(t)$$

and differentiate both sides with respect to t:

$$\text{(19)} \qquad \frac{d}{dt}E[X(t)] = n = 0 \ \ n\frac{d}{dt}p_n(t)$$

$$=-n = 0 \ \ n^2(\lambda+\mu)p_n(t)+\lambda n = 0 \ \ n(n-1)p_{n-}(t)+\mu n = 0 \ \ n(n+1)p_{n+1}(t)$$

where the last equality follows from (16). Setting n-1=n and n+1=n in the second and third sums on the right-hand side (r.h.s.) of (19), respectively, after some straightforward algebra we are led to the following differential equation:

$$\text{(20)} \qquad \frac{d}{dt}E[X(t)] = (\lambda+\mu)E[X(t)]$$

On the other hand from (19) and (20) we obtain the initial condition

$$\text{(21)} \qquad E[X(0)] = j$$

From (20) and (21) we immediately get

$$\text{(22)} \qquad E[X(t) / X(0) = j] = j^*\exp[(\lambda-\mu)t]$$

By a similar procedure an equation for

$$\text{(23)} \qquad \mathrm{Var}\ \{X(t)\} = E[X^2(t)] - \{E[X(t)]\}^2$$

can be derived and solved with the initial condition

$$\mathrm{Var}\ \{X(0)\} = 0$$

We have the following result:

$$\text{(24)} \qquad \mathrm{Var}\ \{X(t)/ X(0) = j\} = \begin{cases} 2j\mu t, & \lambda = \mu \\ \dfrac{j(\lambda+\mu)}{\lambda-\mu}e^{(\lambda-\mu)t}(e^{(\lambda-\mu)t}-1), & \lambda \neq \mu \end{cases}$$

Note that in the case $\lambda \neq \mu$ this variance not only depends on the intrinsic "growth rate" $\lambda - \mu$ but also on the sum $\lambda + \mu$. The sum $\lambda + \mu = R$ defines the noise strength, called demographic noise (Nisbet and Gurney 1982).

Now we turn to the solution of equation (16). We first introduce the monument generating function, $M(\theta, t)$, defined as follows in terms of the p_ns:

$$M(\theta, t) = n = 0 \; e^{\theta n} p_n(t) \tag{25}$$

where p is a dummy variable. Multiplying both sides of equation (16) by $e^{\theta n}$ and summing over n from 0 to ∞, one can easily obtain:

$$\frac{\partial M(\theta,t)}{\partial t} = [\lambda(e^{\theta}-1)+\mu(e^{-\theta}-1)]\frac{\partial M(\theta,t)}{\partial \theta} \tag{26}$$

with the initial condition

$$M(\theta, 0) = e^{\theta j} \tag{27}$$

The general solution of equation (26) can be found by the method of characteristics,

$$M(\theta,t) = \begin{cases} \left[\dfrac{\mu v(\theta,t)-1}{\lambda v(\theta,t)-1}\right]^j & ,\lambda \neq \mu \\ \left[\dfrac{1-(\lambda t-1)(e^{\theta}-1)}{1-\lambda t(e^{\theta}-1)}\right]^j & ,\lambda = \mu \end{cases} \tag{28}$$

where $v(\theta,t) = \frac{(e^{\theta}-1)e^{(\lambda-\mu)t}}{\lambda e^{\theta}-\mu}$

Now we define function $F(s, t)$ as the probability generating function:

$$F(s, t) = n = 0 \; p_n(t)s^n = M(\log(s), t) \tag{29}$$

$$F(s,t) = \begin{cases} \left[\dfrac{\mu(1-\alpha)-(\lambda-\mu\alpha)s}{\mu-\lambda\alpha-\lambda(1-\alpha)s}\right]^j & ,\lambda \neq \mu \\ \left[\dfrac{1-(\lambda t-1)(s-1)}{1-\lambda t(s-1)}\right]^j & ,\lambda = \mu \end{cases} \tag{30}$$

where $\alpha = e^{(\lambda-\mu)t}$

The meaning of $\lambda-\mu$ is intuitive. It is the net rate of population change. To understand the meaning of $\lambda+\mu$ we introduce the diffusion approximation. We rewrite equation (16) as

$$\frac{dp_n(t)}{dt} = -[\lambda(n)+\mu(n)]p_n(t)+\lambda(n-1)p_{n-1}(t)+\mu(n+1)p_{n+1}(t) \tag{31}$$

We use Taylor expansions to approximate the various functions of n±1 in (31). This involves treating n as a continuous variable and reinterpreting $p_n(t)$ as a probability density. Discarding all terms in the Taylor expansions that are of third or higher order, we have the approximate equation of (31),

$$\frac{\partial p_n(t)}{\partial t} = -\frac{\partial}{\partial n}[(\lambda(n)-\mu(n))p_n(t)]+\frac{1}{2}\frac{\partial^2}{\partial n^2}[(\lambda(n)+\mu(n))p_n(t)] \tag{32}$$

With (32) it is easy to show that

$$E(dn) = [\lambda(n) - \mu(n)]dt \tag{33}$$

$$E[(dn)^2] = [\lambda(n)+\mu(n)]dt \tag{34}$$

Equation (34) explains why we call $\lambda+\mu$ the strength of demographic noise.

Now we have established simple closed form expressions for the probability $p_n(t)$ and calculated explicitly mean and variance of the population's size. All this allows us to draw an accurate picture of the population's time evolution. In particular, we can easily get information about the extinction probability of the population by taking the limit of (30) as s goes to 0.

$$p_0(t)=\begin{cases}\left[\dfrac{\mu\left(e^{(\lambda-\mu)t}-1\right)}{\lambda e^{(\lambda-\mu)t}-\mu}\right]^j & ,\lambda\neq\mu\\ \left[\dfrac{\lambda t}{1+\lambda t}\right]^j & ,\lambda=\mu\end{cases} \tag{35}$$

Environmental Stochasticity

Environmental fluctuations inevitably produce fluctuations in population levels. A general question is how species dynamics and interactions translate environmental fluctuations into temporal and spatial patterns of population abundance. Because there is a lack of mathematical machinery available to analyze nonlinear multidimensional stochastic processes, environmental fluctuations are incorporated into a deterministic model in a way that is biologically meaningful yet mathematically tractable. The most common way used in the literature is to add a noise term in a differential or difference equation (May 1973; Capocelli and Ricciardi 1974; Tuckwell 1974; Goel and Richter-Dyn 1974; Turelli 1977). Here we discuss the two most common types of stochastic differential equations: $\frac{dx}{dt} = h(x) + k(x)\Lambda(t)$ and $dx = h(x)dt + k(x)dW$.

Stochastic Differential Equations (1)

$$\frac{dx}{dt} = h(x) + k(x)\Lambda(t)$$

We first deduce the so-called kinetic equation from Markovian property. The kinetic equation is the general form of many stochastic differential equations. Consider a Markov process with a continuum of state values in continuous time. Its transition probability density function (p.d.f.), $f(x, t \mid x_0, t_0)$, satisfies the Chapman-Kolmogrov equation (Bartlett 1966; Ross 1983; Bharucha-Reid 1960):

$$f(x, t \mid x_0, t_0) = \int dy f(x, t \mid y, \tau) f(y, \tau \mid x_0, t_0) \tag{36}$$

with $t>\tau>t_0$ arbitrary instants and $X(t)=x$, $X(\tau)=y$, $X(t_0)=x_0$. Equation (36) is to be looked at as a compatibility relation holding for any Markov process, but it is not sufficient to determine the process' transition p.d.f. To accomplish this task, further assumptions besides the Markov assumption are necessary. First let us rewrite equation (36) in a differential form,

$$f(x, t+\Delta t \mid x_0, t_0) - f(x, t \mid x_0, t_0) = \int dy f(x, t+\Delta t \mid y, t) f(y, t \mid x_0, t_0) - f(x, t \mid x_0, t_0) \tag{37}$$

Let us now consider an arbitrary function R(x) vanishing at the end points of the state space sufficiently rapidly, together with its derivatives of all orders. Multiplying both sides of (37) by R(x)/Δt and integrating over the state space, we obtain:

$$\begin{aligned} &\int dx R(x) \frac{f(x,t+\Delta t|x_0,t_0) - f(x,t|x_0,t_0)}{\Delta t} \\ &= \frac{1}{\Delta t}\int dx R(x) \int dy f(x,t+\Delta t|y,t) f(y,t|x_0,t_0) \\ &- \frac{1}{\Delta t}\int dx R(x) f(x,t|x_0,t_0) \end{aligned} \tag{38}$$

Substituting the Taylor expansion about the point y for R(x) in the first integral on the right-hand side of equation (38):

$$R(x) = R(y) + \sum_{n=1}^{\infty} \frac{d^n R(y)}{dy^n} \frac{(x-y)^n}{n!} \tag{39}$$

and taking the limit as Δt goes to 0, we obtain:

$$\begin{aligned} \int dx R(x) \frac{\partial f}{\partial t} = {} & \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int dy R(y) f(y,t|x_0,t_0) \int dx f(x,t+\Delta t|y,t) \\ & + \sum_{n=1}^{\infty} \frac{1}{n!} \int dy \{ \frac{d^n R(y)}{dy^n} f(y,t|x_0,t_0) \\ & \times \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int dx (x-y)^n f(x,t+\Delta t|y,t) \} \\ & - \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int dx R(x) f(x,t|x_0,t_0) \end{aligned} \tag{40}$$

or:

$$\int dx R(x) \frac{\partial f}{\partial t} = \sum_{n=1}^{\infty} \frac{1}{n!} \int dx \frac{d^n R(x)}{dx^n} f(x,t|x_0,t_0) A_n(x,t) \tag{41}$$

having used the normalization condition:

$$\int dx f(x,t+\Delta t|y,t) = 1 \tag{42}$$

and having set:

$$A_n(x,t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int dy (y-x)^n f(y,t+\Delta t|x,t), \text{ (n=1, 2, ...)} \tag{43}$$

The integration by parts of the r.h.s. of equation (41), in which the vanishing of R(x) and its derivatives at the ends of integration interval is used, shows that:

$$\begin{aligned} & \int dy \frac{d^n R(y)}{dy^n} f(y,t|x_0,t_0) A_n(y,t) \\ & = (-1)^n \int dx R(x) \frac{\partial^n}{\partial x^n} [A_n(x,t) f(x,t|x_0,t_0)] \end{aligned} \tag{44}$$

Equation (41) thus yields:

$$\int dx R(x) \{ \frac{\partial f}{\partial t} - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} [A_n(x,t) f(x,t|x_0,t_0)] \} = 0 \tag{45}$$

Now due to the arbitrariness of the function R(x), the bracketed terms must be identically zero, and we have our desired result:

(46) $$\frac{\partial f(x,t|x_0,t_0)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} [A_n(x,t) f(x,t|x_0,t_0)]$$

This is the kinetic equation holding under the sole assumption that the process under consideration is Markovian. The functions $A_n(x, t)$ defined by (46) are called infinitesimal moments of the process.

Consider a population growth model with a stochastic growth rate due to the environmental fluctuations:

(47) $$r = \bar{r} + \Lambda(t)$$

where $\Lambda(t)$ is a noise term due to the environmental stochasticity and $\bar{r}$ is a deterministic net growth rate. Then we have a simple stochastic differential equation:

(48) $$\frac{dx}{dt} = x_t * \bar{r} + x_t * \Lambda(t)$$

Extending the situation above, we consider a general linear equation of the type

(49) $$\frac{dx}{dt} = h(x) + k(x)\Lambda(t)$$

where h and k are assigned functions and $\Lambda(t)$ is a stochastic process. Clearly, the solution of (49), x(t), is a random function. Its determination cannot be accomplished unless $\Lambda(t)$ is specified.

Let us assume that $\Lambda(t)$ in (49) is a stationary process with a 0 mean and with a rather narrow and peaked correlation function:

(50a) $$E[\Lambda(t)] = g1 = 0$$

(50b) $$E[\Lambda(t_1)\Lambda(t_2)] = g2(t_1, t_2) = g_2(t_2 - t_1)$$

where $g_2(t)$ is appreciably non-zero only in the neighborhood of t=0 with a very sharp maximum at t=0. More generally, for any group of instants $t_1, t_2, \ldots, t_n$ all lying close to each other we set:

(51) $$E[\Lambda(t_1)\Lambda(t_2) \ldots \Lambda(t_n)] = g_n(t_1, t_2, t_n)$$

and, again, assume that the n-th order correlation function g_n has a sharp maximum at $t_1=t_2=...=t_n$, being otherwise effectively 0. Finally, we assume that when $t_1, t_2, ..., t_r$ are proximal to each other, and also when $t_r+1, t_r+2, ..., t_s$ are proximal but far from the group $t_1, t_2, ..., t_r$ and so on, then:

(52)
$$E[\Lambda(t_1) \ldots \Lambda(t_r)\Lambda(t_r+1) \ldots \Lambda(t_s)\Lambda(t_s+1) \ldots \Lambda(t_p) \ldots]$$

$$= E[\Lambda(t_1) \ldots \Lambda(t_r)]E[\Lambda(t_r+1) \ldots \Lambda(t_s)]E[\Lambda(t_s+1) \ldots \Lambda(t_p)] \ldots$$

$$= g_r([t_1 \ldots t_r)g_s(t_r+1 \ldots t_s)g_p(t_s+1 \ldots t_p) \ldots$$

where the functions g_n have already been qualitatively specified.

All these assumptions about the stochastic process $\Lambda(t)$ appearing in equation (49) may look rather artificial at this stage, but the motivation for them will soon be apparent. With this in mind, let us perform a change of variable in (49) by setting:

(53)
$$y = \Phi(x) \quad x=\Phi^{-1}(Y)$$

with

(54)
$$\Phi(x) = \int^{x} \frac{dz}{k(z)}$$

Then, equation (49) is changed into:

(55)
$$\frac{dy}{dt} = \frac{H(y)}{K(y)} + \Lambda(t)$$

upon setting:

(56)
$$H(y) = h[\Phi^{-1}(Y)]$$

$$K(y) = k[\Phi^{-1}(Y)]$$

The advantage of this procedure is that we have constructed a stochastic process y(t) defined by the simpler equation (55) in which $\Lambda(t)$ appears in a purely *additive way.* Due to the above assumptions on $\Lambda(t)$, we now expect y(t), and hence x(t), to be Markovian. Its transition p.d.f. $f_y(y, t/y_0)$ thus satisfies the kinetic equation which we have derived in (46).

(57) $$\frac{\partial fy}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n [B_n f_n]}{\partial y^n}$$

To evaluate the infinitesimal moments B_n we first express the increment of y over a small time interval dt by means of the approximation:

(58) $$dy = y(t+dt) - y(t) \quad \approx \frac{H(y)}{K(y)} dt + \int_{t}^{t+dt} \Lambda(\tau) d\tau$$

where, here and in the following, the value of the process at time t is considered as fixed. Note that equation (58) requires that H(y) and K(y) be smooth; however, the smoothness of the sample paths of Λ(t) is not implied. Taking the expectation of both sides of (58), due to (50), in the limit as $dt \to 0$ we obtain:

(59) $$B_1(y) = \lim_{\partial t \to 0} \left(\frac{1}{\partial t} E[\partial y]\right) = \frac{H(y)}{K(y)}$$

To calculate B_2 we now square both sides of (58) and obtain:

(60) $$(dy)^2 \approx O[(dt)^2] + 2dt \frac{H(y)}{K(y)} \int_{t}^{t+dt} \Lambda(\tau) d\tau + \int_{t}^{t+dt} \Lambda(\tau) d\tau \int_{t}^{t+dt} \Lambda(\theta) d\theta$$

Upon taking the expectations and after dividing by dt, for small dt we are left with:

(61) $$B_2 dt \approx E[(dy)^2] = \int_{t}^{t+dt} \Lambda d\tau \int_{t}^{t+dt} g_2(\tau-\theta) d\theta$$

after making use of equation (50). Using the earlier specified qualitative behavior of g_2, it then follows:

(62) $$B_2 \approx \sigma^2 \text{ with } \sigma^2 = \int_{-\infty}^{\infty} g_2(\mu) du$$

and with the result becoming exact in the limit as $dt \to 0$.

Proceeding along similar lines, it is not difficult to become convinced that due to the assumed properties (50) and (52) of Λ(t) the following relationship holds:

(63) $$dt^* B_n(y) \approx E[(dy)^n] = o(dt) \; (n = 3, 4, \ldots)$$

Equation (57) thus becomes a forward diffusion equation:

(64) $$\frac{\partial f_y}{\partial t} = -\frac{\partial}{\partial y}[B_1 f_y] + \frac{\dot{s}^2 \partial^2 f_y}{2\partial y^2}$$

with $B_1(y)$ and σ^2 given by (59) and (62), respectively. The conclusion is that equation (55) can be thought of as defining a diffusion process y(t) whose drift equals the deterministic part of the r.h.s. of the equation while its infinitesimal variance depends exclusively on the characteristics ($g_2(\tau)$) of the random part of the equation. Furthermore, if we impose that $p\{y(0)=y_0\}=1$, with y_0 being non-random, then the specification of such a diffusion process is unique.

Let us now examine the infinitesimal moments $A_n(x)$ of the Markov process x(t) defined by equation (49). Denoting its transition p.d.f by $f_x(x, t/x_0)$, we know that for small dt we have:

(65) $$A_n(x)dt \approx \int (x'-x)^n f_x(x', dt/x)dx'$$

$$= \int [\Phi^{-1}(y')-x]^n f_x[\Phi^{-1}(y'), dt / x]d[\Phi^{-1}(y')]$$

$$= \int [\Phi^{-1}(y')-x]^n f_y[y', dt/\Phi(x)]dy' \quad (n = I, 2, ...)$$

having made use of the one-to-one transformation (53) between the transition p.d.f.s of the processes x(t) and y(t). Let us now expand $\Phi^{-1}(y')$ as a Taylor series about the point $y=\Phi(x)$:

(66) $$\Phi^{-1}(y') = \Phi^{-1}[\Phi(x)] + \alpha_1(x)[y'-\Phi(x)] + \frac{1}{2}\alpha_2(x)[y'-\Phi(x)]^2$$

$$+ n = 3 \ \frac{a_n(x}{}, n!$$

with $\alpha_n(x) = \frac{d^n F^{-1}(y'}{}, dy^n \mid_{y=\Phi(x)}$

It is easy to see:

(67) $\alpha_1 = k(x)$

$\alpha_2 = k'(x)k(x)$

where k' (x)=dk(x)/dx

Using (66) and (67) and taking the limit $dt \rightarrow 0$, we obtain:

(68) $$A_1(x) = k(x)B_1[\Phi(x)] + \frac{1}{2}k'(x)k(x)B_2[\Phi(x)] + \sum_{n=3}^{\infty}\frac{a_n(x)}{n!}B_n[\Phi(x)]$$

Making use of equations (59) and (61) we thus find:

(69) $$A_1(x) = h(x) + \frac{s^2}{4}\,\frac{dk^2(x}{\quad},dx$$

Using the same procedure, we can have:

(70) $$A_2(x) = \sigma^2k^2(x)$$

$$A_n(x) = 0 \qquad (n=3, 4, \ldots)$$

Thus, we conclude that the Markov process x(t) defined by equation (49) is a diffusion process with drift and infinitesimal variance given by:

(71) $$A_1(x) = h(x) + \frac{s^2}{4}\,\frac{dk^2(x}{\quad},dx$$

$$A_2(x) = \sigma^2k^2(x)$$

Stochastic Differential Equations (2)

$$dx = h(x)dt + k(x)dW$$

In plant communities the effect of environmental fluctuations is cumulative with plants through time. Although the environmental stochastic influence at each time point can be considered as some noise, the cumulated realistic noise could be some other process. This case can be described by the following equation:

(72) $$dx = h(x)dt + k(x)dW$$

where W is some stochastic process such as Brownian motion or Wiener process. Equation (72) is called an Ito equation. This equation can be handled by means of Ito's calculus which differs in several fundamental ways from classic calculus

(Rozovskii 1990). Here we give one of the major results of Ito's calculus: Ito's formula, which can be used to find the solution of the stochastic model (72).

Ito's Formula: Let U: [0, ∞)´R^N be a continuous function with continuous partial derivatives on (0, ∞)´R^N:

(73) $$u_0 = \frac{\partial u}{\partial t}$$

$$u_i = \frac{\partial u}{\partial X_i} \qquad i=1, 2, ..., N$$

$$u_{ij} = \frac{\partial^2 u}{\partial X_i \partial X_j} \qquad i=1, 2, ..., N$$

and let X_1, X_2, ..., X_n be stochastic integrals defined by:

(74) $$X_i(t) = X_i(0) + \int_0^t \varphi_i dW_s + \int_0^t \psi_i d_s$$

then Y defined by:

(75) $$Y(t) = u(t, X_1(t), X_2(t), ..., X_N(t))$$

is also a stochastic integral and its stochastic differential is:

(76) $$dY = u_0 dt + \sum_{i=1}^{N} u_i dX_i + \frac{1}{2}\sum_{i=1}^{N} \sum_{j=1}^{N} u_{ij} dX_i dX_j$$

and the stochastic integral is:

(77) $$Y(t) = Y(0) + {}_0^t\; u_0(s, X_1, X_2, ..., X_N)ds + \sum_{i=1}^{N} {}_0^t\; u_i(s, X_1, X_2, ..., X_N)dXi +$$

$$\frac{1}{2}\sum_{i=1}^{N} \sum_{j=1}^{N} {}_0^t\; u_{ij}(s, X_1, X_2, ..., X_N)dX_i dX_j$$

$$Y(t) = Y(0) + {}_0^t\; u_0 d_s + \sum_{i=1}^{N} {}_0^t\; u_i \varphi_i dW_s + \sum_{i=1}^{N} {}_0^t\; u_i \psi_i d_s +$$

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}{}_{0}^{t}\quad u_{ij}\varphi_i\psi_i d_s$$

Three simple cases of equation (75) are (1) $Y(t)=u(W_t)$, (2) $Y(t)=u(t, W_t)$, and (3) $Y(t)=u(t, X(t))$.

Using Ito's formula, we can easily find the solution of (72). From (76) we can see that the solution x(t) is also a diffusion processes. Its drift and infinitesimal variance are given by:

(78) $B_1(x) = h(x)$

(79) $B_2(x) = \sigma^2 k^2(x)$

To make this discussion more concrete, we consider the following homogeneous unrestricted population growth model with environmental stochasticity. Other situations such as a density dependence model can be solved in a similar way. Taking the deterministic functions h(x) and k(x) in (58) to be rx(t) and $\alpha x(t)$ respectively, we have the simple model:

(80) $dx(t) = rx(t) + \alpha x(t)dB_t$

where r is the intrinsic growth rate and B_t is a Brownian motion. To solve (66), let $Y(t)=u(x(t)=\ln(x(t))$, where $x(t)>0$. Using (62), we obtain the stochastic differential of u(x).

(81) $du(x) = (r-\frac{a^2}{2})dt + \alpha dB_t$

Thus, the solution for the simple population growth model (80) is

(82) $x(t) = x(0)\exp[(r-\frac{a^2}{2})dt + \alpha dB_t]$

The expectation of x(t) is exp(rt)*E[x(0)]. If x(0) is fixed, the expectation is the same as the deterministic model. In the deterministic situation (i.e., x(t) is nonrandom), $x(t)\rightarrow \infty$ as $t\rightarrow \infty$ whenever r is positive. However, in equation (82) x(t) may $\rightarrow \infty$ or 0 even if r is positive. If r is greater than $a^2/2$, $x(t)\rightarrow \infty$ almost surely as $t\rightarrow \infty$. When r is less than $\alpha x(t)$, x(t) goes to zero almost surely as $t\rightarrow \infty$. When r is equal to $\alpha x(t)$, the population growth is then completely controlled by the environmental fluctuations.

Joint Effect of Demographic and Environmental Stochasticity

In the real world, both demographic and environmental stochasticity are unavoidable. Given the difficulties of analyzing each separately, it should come as no surprise that few multispecies analyses incorporate both. To show how to model the combined effect of environmental and demographic noise, let us start with the birth and death process.

We rewrite the birth and death process as:

(83) $$\frac{dp_n(t)}{dt} = -n(\lambda+\mu)p_n(t) + \lambda(n-1)p_{n-1}(t) + \mu(n+1)p_n+1(t)$$

or

$$\frac{dp_n(t)}{dt}, = -(\lambda_n+\mu_n)p_n(t) + \lambda_{n-1}p_{n-1}(t) + \mu_{n+1}p_{n+1}(t)$$

where $\lambda_n = n\lambda$ and $\mu_n = n\mu$. $\lambda+\mu=R$ is the strength of demographic noise.

We suppose that the deterministic dynamics are density dependent, i.e., the logistic growth:

(84) $$dX/dt = Xr(1-X/K)$$

Environmental fluctuations are included by adding a noise term g(X)dW in (84). The corresponding stochastic equation is given as:

(85) $$dX = Xr(X)dt + g(X)dW$$

where W is a standard Wiener process. Different choices for the function g(X) can be motivated. We take:

(86) $$g(X) = \sigma X$$

In this case σ describes the strength of the random fluctuation of the individual growth rate r(X).

As discussed before, (85) is equivalent to a diffusion equation

(87) $$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}[A(x)P] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[B(x)p]$$

where p(x, t) is the probability density that the individual number X shows the value x at time t. The coefficients are given by:

(88) $$B(x) = \sigma^2 x^2$$

(89) $$A(x) = xr(x) = xr(1-x/K)$$

Next we show that the population dynamics with environmental noise, equation (87), can be approximated by the birth and death process (83). We first discretize the right-hand side of (87) by:

(90) $$\frac{d}{dx}B(x) = \frac{B(x+h) - B(x)}{h}$$

(91) $$\frac{d^2}{dx^2}B(x) = \frac{B(x+h) - 2B(x) + B(x-h)}{h^2}$$

with the discretization length h for variable x (h is usually set to 1). The derivatives of A(x) and P(x) are handled in the same way. A simple rearrangement of (87) with (90) and (91) and the definition:

(92) $$p(nh, t) = p_n(t)$$

results in an equation (83) of the birth and death process. The birth and death rates in this case are:

(93) $$\lambda_n = \frac{1}{2}[\frac{B(nh)}{h^2} + \frac{A(nh)}{h}]$$

(94) $$\mu_n = \frac{1}{2}[\frac{B(nh)}{h^2} - \frac{A(nh)}{h}]$$

In the final step we combine both discrete models, i.e., the discrete description of environmental noise by (83) with (93) and (94) at high individual numbers and the model (83) which describes demographic noise at low individual numbers. Therefore, our final stochastic dynamics model, which is valid for both small and large population communities, is the birth-death model (83) with the following birth and death rates:

(95) $$\lambda_n = \frac{1}{2}[\sigma^2 n^2 + Rn + nr(n)]$$

(96) $$\mu_n = \frac{1}{2}[\sigma^2 n^2 + Rn - nr(n)]$$

where Rn is the demographic noise, and r(n) is the density dependent intrinsic growth rate. The simplest form of r(n) is the type of logistic growth rate, i.e., $r(n)=r(1-n/K)$.

A natural community usually has more than one species. The Lotka-Volterra or the Kolmogorov model can model competition among species if there are only two or three species in the community. Using the Lotka-Volterra model or the Kolmogorov model in a highly diverse community may cause chaos because of the feedback. This problem can be solved by introducing the growth rate function r(n) as some other form of the resources. For example, if the total available resource of the community is K, and the used resource at time t, is K_t, then we can define the growth rate of the i-th species as $r^i(n)=r^i*(1-K_t/K)$ or some other form which can express the competition relationship among species.

For the combined model, we want to emphasize the following points. If the density dependence term n/K and the noise term $\sigma^2 n^2$ in equations (95) and (96) are omitted, we get the model with only demographic noise. This model is motivated by simple biological arguments and has been used with small populations by many authors. When the demographic noise term, Rn, is omitted, the resulting model is a diffusion equation, which has been considered by many authors to be a description of environmental noise.

4 Stochastic Estimation and Identification

Quantitative population study has been very popular for several decades and its models have been well developed. These models fall in the categories: dynamic models and stochastic models. In their book, *Modeling Fluctuating Populations*, Nisbet and Gurney (1982) present the most common types of dynamics and stochastic models and the application conditions for each.

One of the basic discrete dynamics models of plant population is the logistic model (Nisbet and Gurney, 1982):

(1) $x_{k+1} = x_k(a - bx_k) - H_k + I_k$, k = 0, 1, ...

where, x_{k+1} and x_k are the population size at time k+1 and k, respectively, h(x,t) is the harvesting, I is immigration or fertilization, and a and b are model parameters.

Consider a more general deterministic system:

(2) $x_{k+1} = f_k(x_k, u_k)$, k = 0, 1, ...

where $x_k \in R^n$ is the state and $u_k \in R^m$ is the input at time k.

One property of equation (2) is that if the current state x_k and the input sequence of the system, u_k, u_{k+1}, ..., u_{k+m} are given, the future states, x_{k+1}, x_{k+2}, ..., x_{k+m+1} can be determined exactly. That is, the deterministic model (2) makes exact prediction of the future population of the plant system.

Under special laboratory conditions and in some isolated environments, population growth may act in a deterministic way. However, most natural populations do not behave as nicely as described by the deterministic model (Hallam 1986). Populations fluctuate about some deterministic trend due to demographic stochasticity and environmental stochasticity (Turelli 1986). In Chapter 3 we discussed how to model stochastic systems.

A stochastic model is able to predict the probability that at a given time, the population will be of a particular size. Given the probability or the conditional probability of future behavior, the expectation of the population size can be determined as it is in the deterministic model. In the real world, however, the probability distributions of most plant populations are rarely known. Because we often do not have descriptive information for the probability distribution, past data must be examined. The availability of observations from the past allows for an alternative method of population modeling: the technique of regression. This technique has become a common tool in estimation and prediction in many areas.

One serious problem with regression models is the assumption of constant coefficients. Because of this assumption, underlying phenomena can never be detected. Also, the reliability of the estimate of regression depends on the sample size. A regression model usually requires a minimal sample size of 30. However, population data that spans 30 or more years is not common. Sometimes we cannot afford to have a large sample size. Moreover, observed data used in regression is assumed to be correct. This is not always the case. In the real world, most natural communities are full of noise. Observed data contains errors such as system noise and measurement error. For some regression models, these errors not only create a large variance, but also produce a biased estimate (Gertner, Cao, and Zhu 1995).

To get a better grip on these problems, we introduce the theory of optimal estimation. An estimator is a process by which information can be extracted from data, i.e., to infer desired information by filtering out the noise from the data. Because the estimator combines the descriptive information and the data information of a system to form an estimate, the estimate usually has a lower variance than a conventional regression estimate. The most common estimator is the Kalman filter (Kalman 1960). The filter describes how to process the measurement data for a given linear system. The theory of optimal application has been successfully applied in a broad range of areas. These areas include signal processing in communications, tracer studies in medicine, statistical image enhancement, estimation of traffic densities, chemical process control, satellite orbit estimation, unclear reactor parameter identification (Gelb 1974), and dendroclimatology (Visser and Molenaar 1988).

Optimal Estimation

An optimal estimator is a computational algorithm that processes measurements in order to deduce a minimum error estimate for the state of a system. It does this by using knowledge of system and measurement dynamics, assumed statistics of system noise and measurement errors, and with initial condition information. The advantages of this type of data processor are that it minimizes the estimation error in a well-defined statistical sense and that it uses all measurement data plus prior knowledge about the system.

There are three types of estimation problems: filtering, smoothing, and prediction (Gelb 1974). When the time at which an estimate is desired occurs at the last measurement point, the problem is referred to as filtering. When the time of interest falls within the span of available measurement data, the problem is termed smoothing. When the time of interest occurs after the last available measurement, the problem is called prediction. In the following sections we will discuss linear filtering and prediction, and two techniques of nonlinear approximation. For convenience, we consider a stochastic system without input. Incorporating a known input is straightforward.

Optimal Linear Filtering

First let us consider a discrete stochastic system without input control:

$$\begin{aligned} x_{k+1} &= A_k x_k + G_k w_k \\ y_k &= C_k x_k + H_k v_k , \end{aligned} \tag{3}$$

where $x_k \in R^n, y_k \in R^p, w_k \in R^g, v_k \in R^h$; A_k, G_k, C_k, and H_k are possible time-varying, known matrices of appropriate dimension, x and y are, respectively, the state space and observation space. The basic random variables $\{x_0, w_0, ..., v_0, ...\}$ are all independent and Gaussian, with $x_0 \sim N(0,\Sigma_0)$, $w_k \sim N(0,Q)$, $v_k \sim N(0,R)$. The covariances are all known. The available information at time k is $z^k = y^k := (y_k, y_{k-1}, ..., y_0)$. The random variable x_k, x_{k+1}, and y^k are jointly Gaussian; denote

$$\begin{aligned} &p_{k|k}(x_k|y^k) \sim N(x_{k|k}, \Sigma_{k|k}), and \\ &p_{k+1|k}(x_{k+1}|y^k) \sim N(x_{k+1|k}, \Sigma_{k+1|k}). \end{aligned} \tag{4}$$

By definition,

$$\begin{aligned} &x_{k|k} := E\{x_k | y^k\}, and \\ &\Sigma_{k|k} := E\{(x_k - x_{k|k})(x_k - x_{k|k})^T | y^k\}. \end{aligned} \tag{5}$$

Similarly,

$$\begin{aligned} &x_{k+1|k} := E\{x_{k+1} | y^k\}, and \\ &\Sigma_{k+1|k} := E\{(x_{k+1} - x_{k+1|k})(x_{k+1} - x_{k+1|k})^T | y^k\}. \end{aligned} \tag{6}$$

To obtain the recursion rule, we take the following steps.

Step 1

From (3),

(7) $x_{k+1|k} = A_k x_{k|k.}$

For convenience denote

(8) $$\begin{aligned} \tilde{x}_{k+1|k} &:= x_{k+1} - x_{k+1|k}, \\ \tilde{x}_{k|k} &:= x_k - x_{k|k}. \end{aligned}$$

From (3) we have

(9) $\tilde{x}_{k+1|k} = A_k \tilde{x}_{k|k} + G_k w_k$

By the independence of $\tilde{x}_{k|k}$ and w_k,

(10) $\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k^T + G_k Q G_k^T$

Step 2

Denote

(11) $$\begin{aligned} y_{k|k-1} &:= E\{y_k | y^{k-1}\}, and \\ \tilde{y}_{k|k-1} &:= y_k - y_{k|k-1} \end{aligned}$$

Since v_k and y^{k-1} are independent, this gives

(12) $$\begin{aligned} y_{k|k-1} &= C_k x_{x|k-1,} \\ \tilde{y}_{k|k-1} &= C_k \tilde{x}_{k|k-1} + H_k v_k. \end{aligned}$$

By the independence of $\tilde{x}_{k|k}$ and v_k

(13) $\Sigma_{k|k-1}^{y} := \mathrm{cov}(\tilde{y}_{k|k-1}) = C_k \Sigma_{k|k-1} C_k^T + H_k R H_k^T$

From (12) and the independence of $x_{k|k-1}$ and $\tilde{x}_{k|k-1}$, we also get

(14) $E x_k y_{k|k-1}^T = \Sigma_{k|k-1} C_k^T$

Step 3

By the innovation principle, we have

$$(15)\qquad x_{k|k} = E\{x_k | y^{k-1}, y_{k|k-1}\}$$

$$= E\{x_k | y^{k-1}\} + (E x_k y^T_{k|k-1})(\Sigma^y_{k|k-1})^{-1} y_{k|k-1}$$

$$= x_{k|k-1} + \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + H_k R H_k^T]^{-1} y_{k|k-1}.$$

and

$$(16)\qquad \Sigma_{k|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + H_k R H_k^T]^{-1} C_k \Sigma_{k|k-1}.$$

Therefore, the conditional density $p_{k|k} \sim N(x_{k|k}, \Sigma_{k|k})$ can be obtained from the following recursion relations.

$$(17)\qquad x_{k+1|k+1} = A_k x_{k|k} + L_{k+1}[y_{k+1} - C_{k+1} A_k x_{k|k}],$$

$$x_{0|0} = L_0 y_{0,}$$

$$\Sigma_{k+1|k+1} = (I - L_{k+1} C_{k+1}) \Sigma_{k+1|k,}$$

$$\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k^T + G_k Q G_k^T$$

$$\Sigma_{0|0} = (I - L_0 C_0) \Sigma_0,$$

here

$$L_k = \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + H_k R H_k^T]^{-1}$$

$$L_0 = \Sigma_0 C_0^T [C_0 \Sigma_0 C_0^T + H_0 R H_0^T]^{-1}$$

The recursion algorithm (17) specifies the transition function of the information state. It is known as the discrete Kalman filter. The matrix L in the recursion is called the Kalman gain matrix (Gelb 1974). It can be shown (e.g., Otter 1978) that the ordinary least squares (OLS) fitting procedure is a special case of the Kalman filter (17).

The transition from the discrete to the continuous formulation of the Kalman filter is straightforward by setting time k and k+1 in the discrete case to time t_k

and t_{k+1}, respectively, and letting $t_{k+1} - t_k = \Delta t \to 0$. First, consider the following continuos stochastic system:

(18) $$\dot{x} = Fx + Gw$$

$$y = Cx + v$$

where w, v are white noise processes with mean 0 and spectral density matrix Q and R, respectively. In this continuous case, v is equivalent to $H_k v_k$ in the discrete case. To make the transition, let us first rewrite the discrete system as a difference equation.

(19) $$x_{t_{k+1}} = A_k x_{t_k} + G_k w t_{k_k}$$

$$\frac{x_{t_{k+1}} - x_{t_k}}{\Delta t} = \frac{(A_k - I)}{\Delta t} x_{t_k} + \frac{G_k}{\Delta t} w_{t_k}$$

From this difference we have the following equivalence, valid in the limit as $t_{k+1} - t_k = \Delta t \to 0$:

$$A_k \to I + F\Delta t$$

In the discrete case we have shown that:

(20) $$\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k^T + G_k Q G_k^T$$

This is now rewritten as:

$$A_k \to I + F\Delta t$$

In the discrete case we have shown that:

(21) $$\Sigma_{k+1|k} = (I + F\Delta t)\Sigma_{k|k}(I + F\Delta t)_k^T + G_k Q G_k^T \Delta t$$

Expansion yields:

(22) $$\Sigma_{k+1|k} = \Sigma_{k|k} + (F\Sigma_{k|k} + \Sigma_{k|k} F^T + G_k Q G_k^T)\Delta t + O(\Delta t^2)$$

From the discrete model we also have:

(23) $$\Sigma_{k|k} = (I - L_k C_k)\Sigma_{k|k-1}$$

Insert (23) into (22) and rearrange the terms we have:

$$\frac{\Sigma_{k+1|k}-\Sigma_{k|k-1}}{\Delta t} = F\Sigma_{k|k-1} + \Sigma_{k|k-1}F^T + G_k Q G_k^T - \frac{1}{\Delta t} L_k C_k \Sigma_{k|k-1} \\ - F L_k C_k \Sigma_{k|k-1} - L_k C_k \Sigma_{k|k-1} F^T + O(\Delta t) \tag{24}$$

Now let us exam the term $L_k / \Delta t$

$$\begin{aligned} \tfrac{1}{\Delta t} L_k &= \tfrac{1}{\Delta t} \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + H_k R H_k^T]^{-1} \\ &= \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T \Delta t + H_k R H_k^T \Delta t]^{-1} \\ &= \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T \Delta t + R]^{-1} \end{aligned} \tag{25}$$

Taking the limit, we get:

$$\lim_{\Delta t \to 0} \tfrac{1}{\Delta t} L_k = \Sigma C^T R^{-1} \tag{26}$$

and,

$$\dot{\Sigma}(t) = F(t)\Sigma(t) + \Sigma(t)F^T(t) + G(t)Q(t)G^T(t) \\ - \Sigma(t)C^T(t)R^{-1}(t)C(t)\Sigma(t) \tag{27}$$

$$\Sigma(0) = \Sigma_0$$

By similar manipulation, we have the state estimate of the continuos system.

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + L(t)[y(t) - C(t)\hat{x}(t)] \tag{28}$$

The matrix gain is:

$$L(t) = \begin{cases} \Sigma(t)C^T(t)R^{-1}(t), & E\left[w(t)v'(\tau)\right] = 0 \\ \left[\Sigma(t)C^T(t) + G(t)\Phi(t)\right]R^{-1}(t), & E\left[w(t)v'(\tau)\right] = \Phi(t)\delta(t-\tau) \end{cases} \tag{29}$$

Optimal Prediction

The prediction model can be derived from the filtering model in a straightforward fashion. The one-step predictor is given by equations (7) and (10).

(30)
$$x_{k+1|k} = A_k x_{k|k.}$$
$$\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k^T + G_k Q G_k^T$$

where $x_{k|k}$ and $\Sigma_{k|k}$ can be calculated from the filtering estimator (17).

Using equation (3), we can derive the n-step predictor.

$$x_{k+n} = A_{k+n-1} x_{k+n-1} + G_{k+n-1} w_{k+n-1}$$

(31)
$$x_{k+n} = \prod_{i=0}^{n-1} A_{k+i} x_k + \sum_{i=0}^{n-1} G_{k+i} w_{k+i}$$

Therefore, the n-step predictor is:

(32)
$$x_{k+n|k} = \prod_{i=0}^{n-1} A_{k+i} x_{k|k}$$
$$\Sigma_{k+n|k} = \prod_{i=0}^{n-1} A_{k+i} \Sigma_{k|k} \left(\prod_{i=0}^{n-1} A_{k+i}\right)^T + \sum_{i=0}^{n-1} G_{k+i} Q G_{k+i}^T$$

Nonlinear Estimation

Now let us consider a more general case described by the nonlinear Stochastic Difference Equation (SDE) with discrete observations.

(33)
$$\dot{x} = f(x,t) + w(t)$$
$$y_k = h_k(x(t_k)) + v_k, \quad k=1, 2, \ldots$$

where f(x,t) and h(x,t) are nonlinear functions, w(t) is Gaussian noise with mean 0 and having spectral density matrix Q(t), and $v_k \sim N(0, R_k)$. One further complicated mode is the form of $\dot{x} = f(x,t) + g(x,t)w(t)$. In this case a theory for estimating x(t) cannot be developed within the traditional framework of mean square stochastic calculus because the r.h.s. of the equation is not integratable in the mean square sense. This difficulty is overcome by formulating the nonlinear filtering problem within the context of Ito calculus.

We first discuss the nonlinear system with the form of (33). There are two widely used linearization techniques: truncated Taylor expansion and statistical approximation (Gelb 1974).

The Taylor expansion method is to write nonlinear function f(x,t) and h(x,t) as a Taylor expansion about the current estimate of the state vector.

$$\text{(34)}\quad \begin{aligned} f(x,t) &= f(\hat{x},t) + \tfrac{\partial f}{\partial x}\big|_{x=\hat{x}}(x-\hat{x}) + \ldots \\ h(x,t) &= h(\hat{x},t) + \tfrac{\partial h}{\partial x}\big|_{x=\hat{x}}(x-\hat{x}) + \ldots \end{aligned}$$

Discarding all the terms with second or higher order, we obtain a linear approximation of f(x,t) and h(x,t) about the current estimate of the state vector. The derivation of the linear filter has been discussed in the previous section.

To obtain an accurate estimate, we need higher-order filters. One method by which the estimate $\hat{x}_k$ can be improved is by repeatedly calculating $\hat{x}_k$, L_k, and $\Sigma_{k|k}$, each time linearizating about the most recent estimate. Another method is to include more terms in the expansions for f(x,t) and h(x,t).

The statistical linearization is to seek a linear approximation for a vector function f(x) of a vector random variable x, having probability density function p(x). The approximation is made by change functions f(x) and h(x) in an approximate linear form.

$$\text{(35)}\quad \begin{aligned} f(x) &\cong E[f(x)] + N_f(x-\hat{x}) \\ h_k(x) &\cong E[h_k(x)] + N_h(x-\hat{x}) \end{aligned}$$

where N_f and N_h are called the function gain matrices. They are estimated by using the technique of minimum mean square error. These estimates are given by:

$$\text{(36)}\quad \begin{aligned} N_f(t) &= [E(fx^T) - \hat{f}\hat{x}^T]\Sigma^{-1}(t) \\ N_h(k) &= \{E[h_k(x_{k|k-1})x^T_{k|k-1}] - \hat{h}_k(x_{k|k-1})\hat{x}^T_{k|k-1}\}\Sigma^T_{k|k-1} \end{aligned}$$

Where $E(fx^T)$, $\hat{f}$, $\hat{x}$ are expectations calculated assumed $x \sim N(\hat{x},\Sigma)$. $E[h_k(x_{k|k-1})x^T_{k|k-1}]$, $\hat{h}_k(x_{k|k-1})$, $\hat{x}_{k|k-1}$ are expectations calculated assumed $x(k) \sim N(x_{k|k-1},\Sigma_{k|k-1})$.

Because nonlinear functions f(x) and h(x) are linearized, the approximate optimal filter for the linearized system now can be found as discussed above. In many instances, the statistical linearization technique has better performance than the Taylor expansion method. However, the decision as to which types of filters should be used in a particular application depends upon their computational complexity and relative performance as observed from realistic computer simulations.

Let us consider a more general nonlinear stochastic system.

(37) $$\dot{x}_t = f(x_t, t) + G(x_t, t)w_t$$

$$y_{t_k} = h(x_{t_k}, t_k) + v_k$$

Now suppose that we generate a reference deterministic trajectory $\overline{x}(t)$, with given $\overline{x}(t_0)$, satisfying:

(38) $$\dot{\overline{x}} = f(\overline{x}(t), t)$$

Define:

(39) $$\delta x_t = x_t - \overline{x}(t),$$

$$\delta y_t = y_t - \overline{y}(t), \text{ and}$$

$$M[t_k, \overline{x}(t_k)] = [\tfrac{\partial h_i(t_k, \overline{x}(t_k))}{\partial x_j}]$$

Linearizing (37) about $\overline{x}(t)$ by a Taylor series expansion, we obtained the linearized discrete system:

(40) $$\delta x_{t_{k+1}} = \Phi[t_{k+1}, t_k, \overline{x}(t_k)]\delta x_{t_k} + w_{t_{k+1}}$$

$$\delta y_{t_k} = M[t_k, \overline{x}(t_k)]\delta x_{t_k} + v_k$$

Where $\Phi[t_{k+1}, t_k, \overline{x}(t_k)]$ is the state transition matrix. That is,

(41) $$d\Phi(t,\tau)/dt = f(t)\Phi(t,\tau), \text{ and}$$

$$\Phi(t,\tau)\Phi(\tau,\xi) = \Phi(t,\xi)$$

The new noise term, $w_{t_{k+1}}$, is given by:

(42) $$w_{t_{k+1}} = \int \Phi(t_{k+1}, \tau)G(\tau)dw_\tau$$

which is a form of Ito's stochastic integral. From the theorem of Ito's stochastic integral, it is easy to see that $\{w_{t_k}\}$ is a white Gaussian sequence with a mean of 0 (Jazwinski 1970), $\{w_{t_{k+1}}\} \sim N[0, Q(k+1)]$, where

$$(43) \qquad Q(k+1)=\int_{t_k}^{t_{k+1}}\Phi(t_{k+1,}\tau)G(\tau)Q(\tau)G^T(\tau)\Phi^T(t_{k+1,}\tau)d\tau$$

Applying the linear filter to the linearized system (40), we obtain the extended Kalman filter for the nonlinear system (37) (Jazwinski 1970).

$$(44) \qquad x_{t_{k+1}|t_k}=x_{t_k|t_k}+\int_{t_k}^{t_{k+1}}f(x_{t|t_k},t)dt.$$

$$\Sigma_{t_{k+1}|t_k}=\Phi[t_{k+1},t_k,x_{t_k|t_k}]\Sigma_{t_k|t_k}\Phi^T[t_{k+1},t_k,x_{t_k|t_k}]+Q(t_{k+1})$$

$$(45) \qquad x_{t_{k+1}|t_{k+1}}=x_{t_{k+1}|t_k}+K[t_{k+1},x_{t_{k+1}|t_k}]\times[y_{t_{k+1}}-h(t_{k+1},x_{t_{k+1}|t_k})]$$

$$\Sigma_{t_{k+1}|t_{k+1}}=\{I-K[t_{k+1},x_{t_{k+1}|t_k}]M[t_{k+1},x_{t_{k+1}|t_k}]\}\Sigma_{t_{k+1}|t_k}\{I-K[t_{k+1},x_{t_{k+1}|t_k}]$$

$$M[t_{k+1},x_{t_{k+1}|t_k}]\}^T+K[t_{k+1},x_{t_{k+1}|t_k}]R(k+1)K^T[t_{k+1},x_{t_{k+1}|t_k}]$$

The Kalman gain is:

$$(46) \qquad K[t_{k+1},x_{t_{k+1}|t_k}]=\Sigma_{t_{k+1}|t_k}M^T[t_{k+1},x_{t_{k+1}|t_k}]\{M[t_{k+1},x_{t_{k+1}|t_k}]$$

$$\times\Sigma_{t_{k+1}|t_k}M^T[t_{k+1},x_{t_{k+1}|t_k}]+R(k+1)\}^{-1}$$

To improve the reference trajectory, we need some iteration algorithms in which the estimate $\eta_{i+1}=x_{t_{k+1}|t_k}+K(t_{k+1},\eta_i)\{y_{k+1}-h(\eta_i,t_{k+1})-M(\eta_i,t_{k+1})[x_{t_{k+1}|t_k}-\eta_i]\}$ can be improved by repeatedly calculating $i=1,\ldots,l$, L_k, and $\Sigma_{k|k}$, each time linearizating about the most recent estimate. The following is the iterated extended Kalman filter with (45) replaced by the iterator (Jazwinski 1970): $\eta_{i+1}=x_{t_{k+1}|t_k}+K(t_{k+1},\eta_i)\{y_{k+1}-h(\eta_i,t_{k+1})-M(\eta_i,t_{k+1})[x_{t_{k+1}|t_k}-\eta_i]\}$, with $i=1,\ldots,l$ and $x_{t_{k+1}|t_{k+1}}=\eta_l$. The iteration starts with $\eta_1=x_{t_{k+1}|t_k}$, and terminates when there is no significant difference between consecutive iterates.

Stochastic Identification

A Kalman filter can yield optimal performance for a linear stochastic system. Consider the discrete stochastic system (3)

$$(47) \qquad x_{k+1}=A_kx_k+B_ku_k+G_kw_k$$

$$y_k=C_kx_k+H_kv_k,$$

The Kalman estimator (conditional mean under given information $x_{k+1|k+1} = A_k x_{k|k} + L_{k+1}[y_{k+1} - C_{k+1} A_k x_{k|k}],$) is given by:

(48) $$x_{k+1|k+1} = A_k x_{k|k} + L_{k+1}[y_{k+1} - C_{k+1} A_k x_{k|k}],$$

The basic assumption of the Kalman estimator is that the descriptions of A, B, G, C, H, Q, R, and Σ_0 [as defined in (3)] are correct. As a practical fact, this is usually impossible.

Estimations of these quantities must be made and improved by the information from observation. Therefore, we need to develop some procedure that provides the best descriptions of A, G, C, H, Q, R, and Σ_0 from observations and other prior knowledge. This problem is called parameter identification.

Consider the following system:

(49) $$y_k = -\sum_{i=1}^{p} a_i y_{k-i} + \sum_{i=1}^{h} b_i u_{k-i} + \sum_{i=0}^{r} c_i w_{k-i}$$

This is an Autoregressive Moving Average Model (ARMAX) model. The first term on the r.h.s. of equation (49) expressing the dependence of the current output on its own past values is the autoregressive, or AR, term. The second term is the external inputs (or control inputs) of the system. The last term, which is called the moving average (MA), is a moving combination of independent random variables v_k.

Using a shift operator, the ARMAX model (49) can be expressed as:

(50) $$A(q^{-1}) y_k = q^{-1} B(q^{-1}) u_k + C(q^{-1}) w_k$$

where q^{-1} is a backward shift operator, i.e., $q^{-1} y_k = y_{k-1}$, and

(51) $$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_n q^{-p} \\ B(q^{-1}) &= b_0 + b_1 q^{-1} + \cdots + b_n q^{-h} \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \cdots + c_n q^{-r} \end{aligned}$$

The stationary condition of an autoregressive process, such as $y_t = a_1 y_{t-1} + \cdots + a_p y_{t-p} + e_t$, is that all roots z_0 of its polynomial, $1 - a_1 z - \cdots - a_p z^p$, have modules $|z_0| > 1$. Also, a moving average process of order r, MA(r), can be written as an infinite AR process if all the roots of its polynomial, $1 - c_1 z - \cdots - c_r z^r$, have modules greater than 1. Such an MA process is called invertible. Without special notice we henceforth assume that the MA term in the ARMAX model is invertible.

Least Square Mean Estimates (LSE)

Suppose we have the data $\{\Phi_0,\cdots,\Phi_{n-1},y_1,\cdots,y_n\}$ and we believe that the following model fits the data.

$$y_{k+1}=\Phi_k^T\theta+w_{k+1} \tag{52}$$

where w_k is the error term with a mean of 0.

Let F_n be a s-field generated by the data $\{\Phi_0,\cdots,\Phi_{n-1},y_1,\cdots,y_n\}$. Then, {$w_k$, k £ n} is F_n-measurable because w_k can be deduced from y_k and F_k through the relation $w_{k+1}=y_{k+1}-\Phi_k^T\theta$.

We also assume that {w_k} is a Martingale difference sequence with respect to the increasing sequence of s-field {$\Im_k$} which satisfies

$$E[w_k^2|\mathcal{F}_{k-1}]=\sigma^2,\ \text{a.s., for all k}$$

and

$$\sup_k E[|w_k|^\alpha|\mathcal{F}_{t-1}]<\infty,\ \text{a.s., for some } \alpha>2.$$

The least square method is to choose θ to minimize

$$V_n(\theta)=\sum_{k=0}^{n-1}(y_{k+1}-\Phi_k^T\theta)^2 \tag{53}$$

By setting $\frac{\partial V_n(\theta)}{\partial\theta}=0$, we obtain the LSE of θ

$$\hat{\theta}_n=(\sum_{k=0}^{n-1}\Phi_k\Phi_k^T)^{-1}\sum_{k=0}^{n-1}\Phi_k y_{k+1} \tag{54}$$

Suppose one more datum (Φ_n,y_{n+1}) becomes available; then we can obtain the new LSE $\hat{\theta}_{n+1}$ from the old LSE $\hat{\theta}_n$.

$$\begin{aligned}\hat{\theta}_{n+1}&=\hat{\theta}_n+R_n^{-1}\Phi_n(y_{n+1}-\Phi_n^T\hat{\theta}_n),\\ R_{n+1}&=R_n+\Phi_n\Phi_n^T.\end{aligned} \tag{55}$$

Where $R_n = \sum_{k=0}^{n} \Phi_k \Phi_k^T$. To verify (55), we substitute the equation

$\hat{\theta}_n = (\sum_{k=0}^{n-1} \Phi_k \Phi_k^T)^{-1} \sum_{k=0}^{n-1} \Phi_k y_{k+1}$ by $\sum_{k=0}^{n-1} \Phi_k y_{k+1} = R_{n-1}\hat{\theta}_n$, and write $R_{n-1} = R_n - \Phi_n \Phi_n^T$

Let $P_n = R_n^{-1}$. We obtain the LSE recursions.

$$\begin{aligned} \hat{\theta}_{n+1} &= \hat{\theta}_n + \tfrac{P_{n-1}\Phi_n}{1+\Phi_n^T P_{n-1}\Phi_n}(y_{n-1} - \Phi_n^T \hat{\theta}_n), \\ P_n &= P_{n-1} - \tfrac{P_{n-1}\Phi_n \Phi_n^T P_{n-1}}{1+\Phi_n^T P_{n-1}\Phi_n}(y_{n-1} - \Phi_n^T \hat{\theta}_n). \end{aligned} \tag{56}$$

For convenience, we define $\tilde{\theta}_k = \hat{\theta}_k - \theta_k$, and $\hat{w}_{k+1} = y_{k+1} - \Phi_k^T \hat{\theta}_k$.

Theorem (LSE convergence)

Consider the LSE algorithm defined by (55) and applied to model (52), then

(i) $\sum_{k=0}^{n} (\hat{w}_k - w_k)^2 = O(\log Tr(R_{n-1}))$ a.s.

(ii) $\|\tilde{\theta}_n\|^2 = O(\frac{\log Tr(R_{n-1})}{\lambda_{\min}(R_{n-1})})$ a.s.

(iii) $\lim_{n\to\infty} \frac{1}{S_{n-1}^i} \sum_{k=0}^{n-1} \Phi_k^i w_{k+1} = 0$ a.s., here i=1, ..., 2p+2,

Moreover, if $\lim_{n\to\infty} \lambda_{\min}(R_n) = +\infty$ and $\frac{R_n}{Tr(R_n)} \geq \varepsilon I$ for all large n, then

$\lim_{n\to\infty} \hat{\theta}_n = \theta$ a.s.

Proof

We will not prove results (i) to (ii) here. They can be proved by using Martingale convergence theorem and Kronecker's lemma (i.e., two real valued sequences $\{x_k\}$ and $\{r_k\}$ satisfying $r_k > 0$, $\lim_{k\to\infty} r_k = \infty$, $\sum_{k=1}^{\infty} \frac{x_k}{r_k} < \infty$, then $\lim_{N\to\infty} \frac{1}{r_N} \sum_{k=1}^{N} x_k = 0$). We only give the proof of (iv), the convergence of the estimate.

Since $S^i_{n-1} = 1 + \sum_{k=0}^{n-1} (\Phi^i_k)^2 \le 1 + Tr \sum_{k=0}^{n-1} \Phi_k \Phi_k^T$, from result (iii) we have

$$\lim_{n\to\infty} \frac{1}{Tr \sum_{k=0}^{n-1} \Phi_k \Phi_k^T} \sum_{k=0}^{n-1} \Phi^i_k w_{k+1} = 0$$

and

$$\hat{\theta}_n = \theta + P_{n-1} \sum_{k=0}^{n-1} \Phi_k w_{k+1}$$

$$\hat{\theta}_n = \theta + \left[\frac{1}{Tr \sum_{k=0}^{n-1} \Phi_k \Phi_k^T} \sum_{k=0}^{n-1} \Phi_k \Phi_k^T\right]^{-1} \times \left[\frac{1}{Tr \sum_{k=0}^{n-1} \Phi_k \Phi_k^T} \sum_{k=0}^{n-1} \Phi_k w_{k+1}\right]$$

Taking the limit of the equation above, we obtain result (iv).

Suppose we have the data $\{\Phi_0, \cdots, \Phi_{n-1}, y_1, \cdots, y_n\}$ at time n. The one-step prediction made by $\hat{\theta}_n$ is

$$\hat{y}_{n+1} = \Phi_n^T \hat{\theta}_n$$

Denote θ_i is the i-th element of θ, then the feedback law is explicitly given by

$$\begin{aligned} u_n = \frac{-1}{\hat{\theta}_n(q+1)} [& \hat{\theta}_n(1) y_n + \cdots + \hat{\theta}_n(q) y_{n-q+1} \\ & + \hat{\theta}_n(q+2) u_{n-1} + \cdots + \hat{\theta}_n(q+h) u_{n-h+1} \\ & - \hat{y}_{n+1} - \hat{\theta}_n(q+h+1) \hat{y}_n - \cdots - \hat{\theta}_n(q+h+r) \hat{y}_{n-r+1} \end{aligned} \tag{57}$$

Similar results can be obtained for stochastic gradient (SG) algorithm. The SG algorithm is regarded as a simplification of the LSE recursions since the scalar gain in the SG algorithm is only the trace of the matrix gain in the LSE algorithm.

Conditional Mean Estimates

Let us rewrite the ARMAX model (49) as

$$y_{k+1} = \Phi_k^T \theta^0 + w_{k+1}$$

where $\Phi_k = (y_t, \cdots, y_{t-p}, u_t, \cdots, u_{t-h}, w_t, \cdots, w_{t-r})^T$, $\theta^0 = (a_1, \cdots, a_p, b_1, \cdots, b_h, c_1, \cdots, c_r)^T$. The true parameter θ^0 is unknown and we want to identify it.

Theorem (conditional mean estimate)

Assume that $\theta^0 \sim N(\overline{\theta},\ \Sigma)$, {yk} is determined by $y_{k+1} = \Phi_k^T \theta^0 + w_{k+1}$, $\{w_k\}$ is identical independent distribution (iid) with $w_k \sim N(0, \sigma^2)$, and $\{y_s, \Phi_s, s \leq k\}$ is independent of $\{w_s, s \geq k+1\}$. Then $\hat{\theta}_n = E\{\theta^0 | \Phi_0, \cdots, \Phi_{n-1}, y_1, \cdots, y_n\}$ minimizes $E\{\|\theta^0 - \theta\|^2 | \Phi_0, \cdots, \Phi_{n-1}, y_1, \cdots, y_n\}$, and $\hat{\theta}_n$ and $P_n := E\{(\theta^0 - \hat{\theta}_n)(\theta^0 - \hat{\theta}_n)^T | \Phi_0, \cdots, \Phi_{n-1}, y_1, \cdots, y_n\}$ satisfy the recursions

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \tfrac{P_{n-1}\Phi_n}{1+\Phi_n^T P_{n-1}\Phi_n}(y_{n-1} - \Phi_n^T \hat{\theta}_n),$$

$$P_n = P_{n-1} - \tfrac{P_{n-1}\Phi_n\Phi_n^T P_{n-1}}{1+\Phi_n^T P_{n-1}\Phi_n}(y_{n-1} - \Phi_n^T \hat{\theta}_n). \tag{58}$$

This theorem says if $\{w_k\}$ is Gaussian white noise, the recursions for the conditional mean of θ^0 under Bayesian formulation coincides with the LSE.

Proof

Let $V(\theta) = E\{\|\theta^0 - \theta\|^2 | \Phi_0, \cdots, \Phi_{n-1}, y_1, \cdots, y_n\}$

$$\min_\theta \{V(\theta)\} \Rightarrow \tfrac{\partial J(\theta)}{\partial \theta} = 0$$

$$\Rightarrow 2E\{|\theta^0 - \theta| \ \ |\Phi_0, \cdots, \Phi_{n-1}, y_1, \cdots, y_n\} = 0$$

$$\Rightarrow \hat{\theta} = E\{\theta^0 | \Phi_0, \cdots, \Phi_{n-1}, y_1, \cdots, y_n\}$$

To prove the recursion (58), we rewrite the ARMAX model as a state space representation.

$$\theta_{n+1} = \theta_n, \quad \theta_0 \sim N(\overline{\theta},\ \Sigma),$$

$$y_{k+1} = \Phi_k^T \theta^0 + w_{k+1}$$

Then we can use the Kalman filter as discussed in the previous section

$$x_{k+1|k+1} = A_k x_{k|k} + L_{k+1}[y_{k+1} - C_{k+1} A_k x_{k|k}],$$

$$\Sigma_{k+1|k+1} = (I - L_{k+1} C_{k+1}) \Sigma_{k+1|k},$$

$$\Sigma_{k+1|k+1} = (I - L_{k+1} C_{k+1}) \Sigma_{k+1|k},$$

$$L_k = \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + H_k R H_k^T]^{-1}$$

and make the following exchanges

$$x_{k|k} \Leftrightarrow \hat{\theta}_k$$

$$\Sigma_{k|k} \Leftrightarrow P_k$$

$$A_k \Leftrightarrow I, \;\; G_k \Leftrightarrow 0, C_k \Leftrightarrow \Phi_k^T, \;\; H_k \Leftrightarrow I$$

We obtain the recursion (58).

5 STELLA II Modeling Process

In the previous chapters, we have chosen and defined theoretical statistical measures of plant diversity and developed theoretical models for plant communities. The stochastic dynamics models developed contain two main components: deterministic process and stochastic process. The deterministic model controls the biological dynamics of the plant communities, while the stochastic process simulates the biological and ecological fluctuations. When we have a good understanding of the dynamics and stochastic behaviors of the plant community we are studying, the stochastic dynamics model is able to make respectable predictions for the community. Once we have chosen to use a stochastic model, there remains the task of selecting the mathematical approach to follow in its analysis. Figure 1 illustrates the sequence of decisions involved in choosing an appropriate model. The first decision is the specification of birth and death probabilities. The second is to determine the size of fluctuation due to demographic variation. Finally, a choice of representation of environmental variation must be made. Following the flow of the diagram, we can select one of the mathematical models discussed in Chapter 3 as our simulation model.

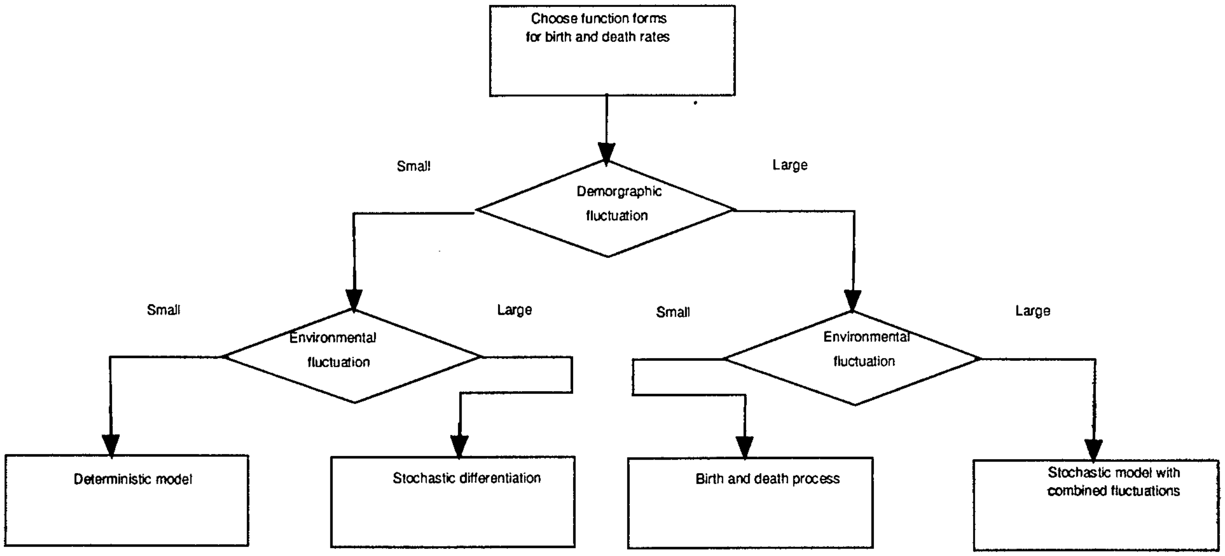


Figure 1. Sequence of choosing a stochastic model.

As we pointed out before, a stochastic model is able to predict the probability that the population will be of a particular size. Given the probability or the conditional probability of future behavior, an estimate of the population can be determined as in the deterministic model. In the real world, however, we usually do not know the probability distribution, and we do not know the details of population dynamics. Instead, we must look at past data. In this case, estimation models discussed in Chapter 4 will be our choice of simulation models.

U.S. military installations across the country cover thousands of plant community types. With the development of the U.S. Army Land Condition Trend Analysis (LCTA) program, Army-wide inventory plots have been established and field data have been collected on many Army installations since 1989. For this reason, we choose the estimation model presented in Chapter 4 as our major STELLA simulation model. (STELLA is a software modeling tool.) The dynamics model is used only for testing purposes.

Consider the discrete stochastic system presented in Chapter 4.

$$(1)\quad \begin{aligned} x_{k+1} &= A_k x_k + G_k w_k \\ y_k &= C_k x_k + H_k v_k, \end{aligned}$$

where $x_k \in R^n, y_k \in R^p, w_k \in R^g, v_k \in R^h$; A_k, G_k, C_k, and H_k are possible time-varying, known matrices of appropriate dimension, x and y are, respectively, the state space and observation space. The basic random variables $\{x_0, w_0, .., v_0, ...\}$ are all independent and Gaussian, with $x_0 \sim N(0, \Sigma_0)$, $w_k \sim N(0, Q)$, $v_k \sim N(0, R)$. The covariances are all known. The available information at time k is $z^k = y^k := (y_k, y_{k-1}, \ldots, y_0)$. The random variable x_k, x_{k+1}, and y^k are jointly Gaussian. As derived in Chapter 4, the recursion scheme was obtained by Kalman filtering:

$$(2)\quad x_{k+1|k+1} = A_k x_{k|k} + L_{k+1}[y_{k+1} - C_{k+1} A_k x_{k|k}],$$

$$x_{0|0} = L_0 y_{0,}$$

$$\Sigma_{k+1|k+1} = (I - L_{k+1} C_{k+1}) \Sigma_{k+1|k,}$$

$$\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k^T + G_k Q G_k^T$$

$$S_{0|0} = (I - L_0 C_0) S_{0,}$$

Here

$$L_k = \Sigma_{k|k\ 1} C_k^T [C_k \Sigma_{k|k\ 1} C_k^T + H_k R H_k^T]^{\ 1}$$

$$L_0 = \Sigma_0 C_0^T [C_0 \Sigma_0 C_0^T + H_0 R H_0^T]^{-1}$$

This recursion specifies the transition function of the information state. It is known as the discrete Kalman filter. The matrix L in the recursion is called the Kalman gain matrix. The following STELLA model uses this recursion scheme to model the changes of the biodiversity and population of plant communities.

The STELLA Model

In this study, the STELLA model is built based on the recursion scheme (2). Figure 2 is a STELLA map, which shows the four basic parts of our STELLA model: global variables, local variables, the Kalman filter, and the outputs. A variable is called global if it is determined by the community type and is independent of the species type, such as the Shannon index, total number of species, total population, total number of observations, error limit, natural input, and human input (these are discussed in the following section). A local variable varies with the type of species. Local variables include species abundance, species tolerance to environmental changes, and model parameters. The model of the Kalman filter is based on species. Although the structure of the model is the same for all species, the inputs and outputs are different for different species. The outputs include all the major results of the simulation both in table and graphic format.

Figure 3 is a STELLA diagram that shows the structure of the estimation model of the Kalman filter. The parameters and components in the model are defined in the recursion scheme (2). This model simulates the changes of the population of a single species. A multispecies community needs multiple copies of this model in which the structure is the same but the parameters are set differently.

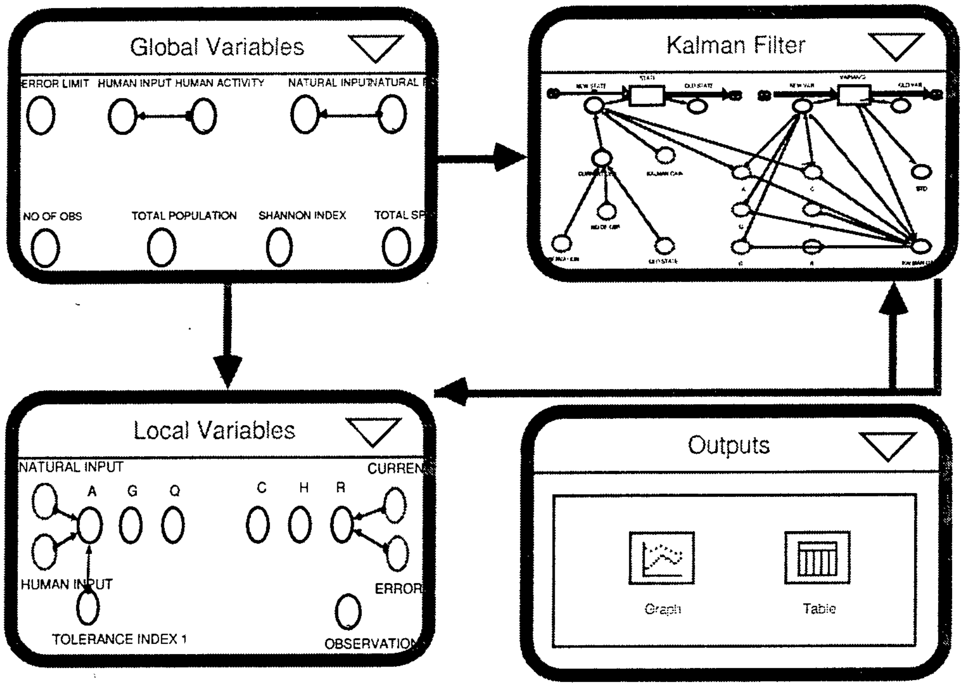


Figure 2. The STELLA map of flow structure.

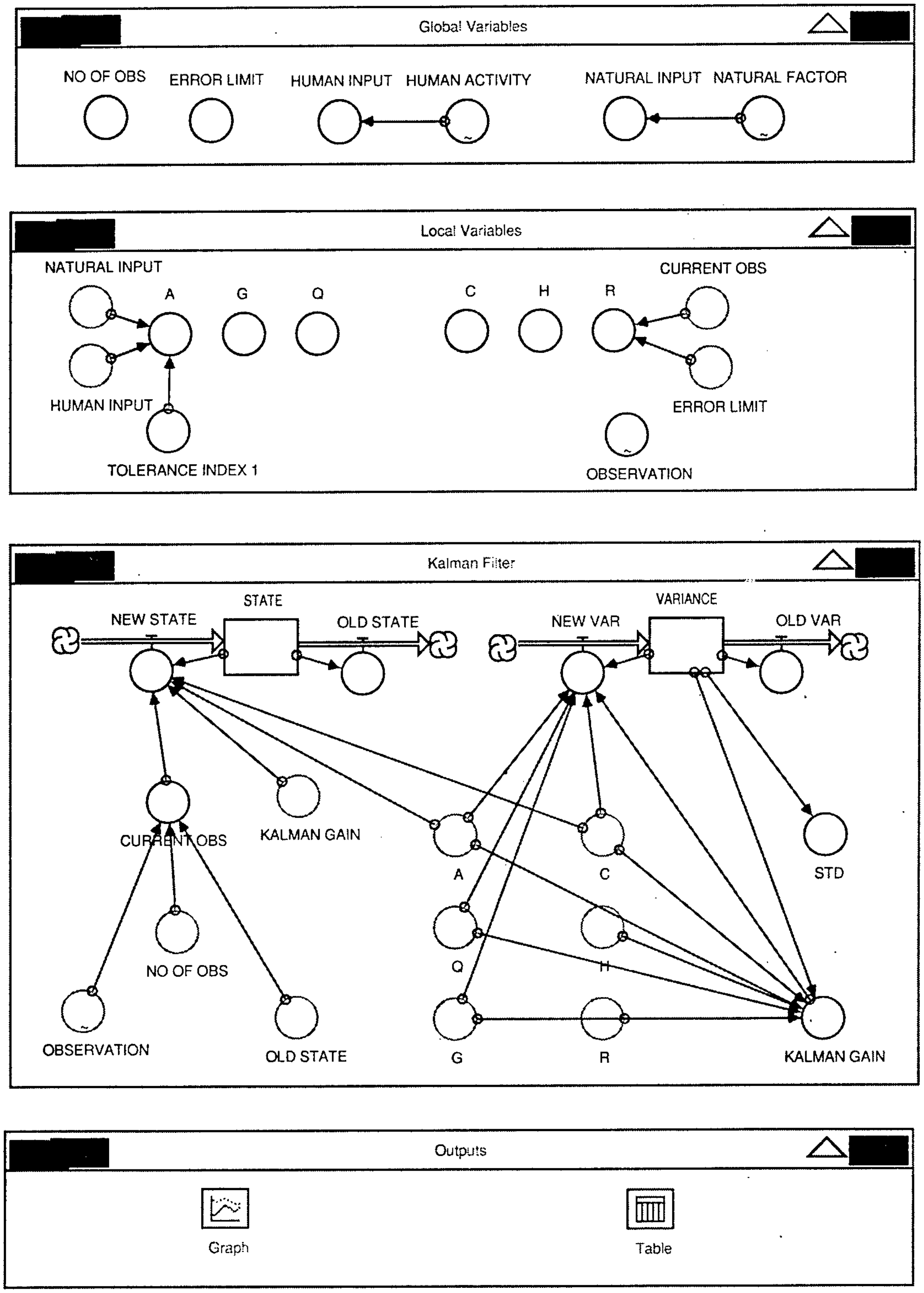


Figure 3. The STELLA structure of the Kalman filter.

Setting of the STELLA model

Global Variables

Number of observations: is the number of repeated times of observation.

Error limit: is the tolerance of inventory errors that include sampling and non-sampling errors.

Human input: is the quantified human activities, including the destruction and improvement activities done to the community.

Natural input: is scaled environmental variation and natural catastrophe. The growth of a plant population is determined not only by its intrinsic reproduction and survival capabilities but also by its surrounding climatic, topographic, and geologic conditions. Some of these environmental factors have little variation over time, such as topography and soil, while others, such as temperature and precipitation, change daily. Some of these changes are important to the population growth, and some of them are not. We choose those factors that are important to the community and change over a relatively short period of time. Catastrophes can be thought of as the extremes of environmental variation. These are events that affect either reproduction or survival. Catastrophes include habitat destruction, flood, fire, disease, drought, storm, etc. We may be able to define the impacts of these catastrophes on the community of interest by examining the historical records of catastrophes and population changes of the community.

Total population: is the sum of the populations of all species in the community.

Total number of species: is the total number of species appearing in the community during the simulation.

Shannon index: is defined as $-\sum_{i=1}^{S} p_i \log(p_i)$, where S is the total number of species and p_i is relative population of the i-th species. The Shannon index is determined by both the richness (number of species of the community) and the evenness (relative abundance of each species). Given the total number of species, the Shannon index reaches its maximum when all species have the same abundance. That is, Max (H) = $\log(S)$. Comparing the Max (H) and the true value of H, we know the level of the diversity of the community.

We also calculate the population sizes of the most abundant species and their variances. In most natural communities, only a few species contribute 80 percent or more to the total population. These abundant species usually determine the quality and function the community reaches.

Local Variables

Model parameters: model parameters A, G, Q, C, H, and R are defined in the schema (20). In our study, parameters G, C, and H are set to be 1. A is the net growth rate. When a community is in a stable stage, parameter A should be 1. Thus A is initialized to be 1. Q is the system noise term. It is usually unknown. As we proved in Chapter 4, Q does not affect the trend of estimates. Even when Q is initialized large, it drops very fast as more and more data enter into the model. R is the combined inventory error, which is determined by the population sized and the error limit.

Species tolerance: the tolerance of a species is the capability of the species to resist environmental changes. It is one of the intrinsic characters of the species. Different species have different sensitivities to changes in their environment. Thus, the tolerance index is determined by our understanding of the species.

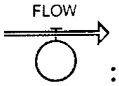
Observation: the observation is a set of time-series data collected from the field. The size of the data set affects the quality of the simulation results. The more data we have, the more actuate the results we obtain from the simulation.

Kalman filter

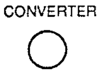
Stocks, Flows, Converters, and Connectors are the generic building blocks of the STELLA language. Stocks, flows, or converters represent variables and parameters. Connectors link related variables or parameters. This section provides a brief description of these building blocks, and shows how the blocks are used to build the Kalman filter (2). STELLA manual (STELLA II 1994) provides details of modeling in STELLA.

STOCK

□: In our model, stocks act as buffers within the system of plant populations. They build up or decline whenever their associated rates of inflow and outflow are out of balance with one another. This buffering property of stocks leads to dynamic simulation of the system. There are two stock variables in the Kalman filter: population and its variance. The population is the optimal estimate of the true population of a given species. The variance is the minimized variance of the population. These two stock variables are calculated by the Kalman filter (2).



: In STELLA, flows consist of a pipe (or conduit), flow regulator, and arrow. Things flow through the conduit in the direction indicated by the arrow. The specific volume of the flow is calculated by the algebraic expression in the regulator. In the Kalman filter, there are two types of flows: flow-in and flow-out. A flow-in and flow-out connect each of the stock variables in our model. When a new observation comes into the model, the regulator in the flow-in calculates the new population size based on the Kalman filter, and adds the result into the stock. Meanwhile, the flow-out removes the old population from the stock. Therefore, the population size stored in the stock is updated as the recursion scheme (2) shows.



: Circles in STELLA represent converters; they are the containers for all types of information or material quantities. As their name implies, converters transform inputs into outputs based on the expressions in the circles. Unlike stocks, converters do not accumulate flows and have no memory. In our model, converters represent all of the global and local variables.

Outputs

The outputs of the model include total population size of the community, total number of species, the Shannon index, and the populations of the five most abundant species of the community and their standard deviations. These outputs are presented in table and graphic formats.

6 A Case Study: White Sands Installation

White Sands Missile Range was selected as the case study. The White Sands installation is located in the desert in New Mexico. Although the species composition here is not complicated, White Sands has a wide range of land covers, from pure sand with little or no vegetative cover, to highly dense plant communities. A single model may not be able to model the whole installation. Therefore, we first classify the plots of White Sands into different plant community types based on their species compositions. Then, we simulate the population and diversity of each type of plant community based on data collected from the field.

Classification of Plant Communities

A plant community is an interacting unit of all the populations of species within a prescribed region. Different communities have different species composition and present different patterns. A central goal of plant ecology and biology is to understand and model the processes and mechanisms that cause the patterns we see. This is no easy task. Any given habitat may contain from a few plant species to hundreds of different species. In a large area, especially a tropical or subtropical area, species composition is usually very complicated and vegetation patterns are complex. No single model can produce a satisfactory description for the population dynamics in all types of plant communities. A model usually works well within a certain type of community. Therefore, to understand the vegetation changes in a given area, the first task we have to accomplish is to classify that community.

In most sampling designs, vegetation types or land cover types are identified based on satellite imagery, soil types, and vegetation information. Samples are then proportionally assigned to the land cover categories classified from satellite imagery. For a variety of reasons, classifications based on satellite imagery do not reflect the natural distribution of plants and land covers. The preliminary classification based on satellite imagery should be modified by observation information.

There are many different approaches to and kinds of community classification. The two basic methods are the ecological classification framework and numerical classification. The ecological classification framework combines the physical and biological factors of a given region. Usually, dominant species of a plant community are used to represent the ecological function of the community. Thus, a direct method for ecological classification is to put samples into groups of dominant-types according their species composition. That is, samples with similar major species are classified into the same group. This method is direct and accurate when all samples are dominated by a few species. However, when the sample size is very large and species composition is very complicated, the direct method may give a large number of dominant types. If the number of dominant types is limited, it is very hard, even for an experienced expert, to tell which group a sample should go into. Moreover, this ecological method is more or less subjective. Different people may have different classification schemes for the same data set.

In the past three or four decades, there has been an increasing tendency to use numerical methods. This has largely been due to the greater objectivity of these methods. A common method is cluster analysis. Clustering is a method of finding groups in data. But the direct use of species abundance data in clustering causes the problem of putting samples with similar species composition into different groups. For example, two samples (I and II) both have species A, B, C, and D. Sample I has the species abundance 100, 60, 40, and 10 of species A, B, C, and D, respectively. Sample II has the species abundance 60, 30, 20, and 5 of species A, B, C, and D, respectively. Although these two samples have similar species composition they may be placed into different groups by cluster analysis because the absolute difference of species composition between these two samples is large.

Instead of using the direct abundance, we standardize the abundance data by quadrate in this study. Then, we apply usual clustering methods to the standardized data. This method overcomes the disadvantages of using direct abundance because the abundance of species is standardized to a scale of unit one for all samples. The clustering programming is written in the SAS statistical analysis program. We use three different clustering methods: average, centroid, and Ward's minimum variance, and choose the best clustering scheme after comparing the clustering results obtained by these three methods.

Standardization of Data by Quadrat

There are many different approaches to standardizing data onto a scale of one. We used quadratic standardization (scaling) in this study. The reason for using quadratic scaling is that this method of scaling puts appropriate weights on abundant species. Relative abundance, which is defined as the i-th species abundance divided by the total population, treats all the species as having the same importance in the classification. This is not what we expect in the classification of a plant community. In community classification, abundant species are usually considered to be more important than other species. On the contrary, cubic or higher order scaling puts too much weight on the abundant species, and other species may have little or no impact on the classification.

To understand quadratic scaling, we first define the length of a vector. The length of a vector is defined as the square root of the sum of the squares of all its elements. That is,

$$L = \sqrt{\sum Q_i^2} \tag{1}$$

where Q_i is the i-th element of vector Q.

Let Q be the abundance vector of a community. Q_i is the abundance of i-th species. Then, the standardized abundance of the i-th species, which is also called the importance index of the i-th species, is defined by:

$$Q_i' = \frac{1}{L^2} Q_i^2 \tag{2}$$

Two properties of the importance index are: (1) Q' is a unit vector; that is, Q_i' ranges from 0 to 1, and $\sum Q_i' = 1$; (2) Q_i' is determined by the relative abundance rather than absolute abundance.

Clustering Method

The SAS clustering procedure, CLUSTER, is used to find groups of observations with coordinate data (species importance index). To obtain a better clustering result, we use three clustering methods: average linkage, centroid method, and Ward's minimum-variance method.

The following notation is used, with lowercase symbols generally pertaining to observations and uppercase symbols to clusters:

n = number of samples (observations);
v = number of variables;
G = number of clusters at any given level of the hierarchy;
x_k = i-th observation;
C_K = K-th cluster;
N_k = number of observations in C_K;
$\bar{x}_k$ = mean vector for C_K;

$\|x\|$ = Euclidean length of the vector x;
$T = \sum_{i=1}^{n} \|x_i - \overline{x}\|^2$;
$W_k = \sum_{i\in C_k} \|x_i - \overline{x_k}\|^2$;
$PG = \sum w_J$;
$B_{KL} = W_M - W_K - W_L$ if $C_M = C_k \cup C_L$;
D(x, y) = distance between vectors x and y;
D_{KL} = distance between clusters C_K and C_L.

The distance between two clusters can be defined either directly or by combinatorial. That is, by an equation for updating a distance matrix when two clusters are joined. In all combinatorial formulas below, it is assumed that clusters C_K and C_L are merged to form C_M, and the formula gives the distance between the new cluster C_M and any other cluster C_J.

Average Linkage

In the average linkage, the distance between two clusters is defined by:

$$(3) \qquad D_{KL} = \sum_{i\in C_K} \sum_{j\in C_L} d(x_i, x_j) / (N_K N_L).$$

if d(x,y) = $\|x - y\|^2$ then:

$$(4) \qquad D_{KL} = \|\overline{x_K} - \overline{x_L}\|^2 + W_K / N_K + W_L / N_L.$$

The combination formula is:

$$(5) \qquad D_{JM} = (N_k D_{JK} + N_L D_{JL}) / N_M.$$

In the average linkage, the distance between two clusters is the average distance between pairs of observations, one in each cluster. Average linkage tends to join clusters with small variances but is slightly biased toward producing clusters with the same variance.

Centroid Method

In the centroid method, the distance between two clusters is defined by:

$$(6) \qquad D_{KL} = \|\overline{x_K} - \overline{x_L}\|^2.$$

If $d(x,y) = \|x - y\|^2$ then the combination formula is:

(7) $D_{JM} = (N_k D_{JK} + N_L D_{JL}) / N_M - N_K N_L D_{KL} / N_M^2$.

In the centroid method the distance between two clusters is defined as the Euclidean distance between their centroid or means. The centroid method is more robust to outliers than most other hierarchical methods.

Ward's Minimum-variance Method

In Ward's minimum-variance method, the distance between two clusters is the analysis of variance (ANOVA) sum of squares between two clusters added up over all the variables. At each generation, the within-cluster sum of squares is minimized over all partitions obtainable by merging two clusters from the previous generation. d(x, y), D_{KL} and the combinatorial formulas are, respectively, defined as:

(8) $d(x, y) = | | x\text{-}y | |^2/2$

(9) $D_{KL} = | | \bar{x}_K - \bar{x}_L | |^2 / (1/N_K + 1/N_L)$

(10) $D_{JM} = ((N_J + N_K)D_{JK} + (N_J + N_L)D_{JL} - N_J D_{KL}) / (N_J + N_M)$

Classification Results

The clustering results were obtained by the three clustering methods: average linkage, centroid method, and Ward's minimum-variance method. Tables 1 and 2 summarize the results of cluster analysis for years 1989 and 1992, respectively. From Tables 1 and 2 we can see Ward's minimum variance method gives the best match of plots among all the three methods and a stable classification scheme between years 1989 and 1992. Thus we choose Ward's clustering scheme as the frame scheme and adjusted it by the clustering results from average linkage and centroid method. We suggest 18 vegetation types for White Sands (17 clusters plus 1 type of bare soil). The vegetation types and their major species are listed in Table 3.

Table 1. Comparison of cluster analysis results for 1989.

Cluster	Average Linkage	Centroid Hierarchical	Ward's Minimum Variance
1	31, 33, 35, 36, 147, 150, 175	31, 33, 35, 36, 147, 150, 175	31, 33, 35, 36, 147, 150, 175
2	1, 2, 10, 11, 12, 13, 14, 24, 29, 30, 37, 41, 42, 54, 59, 73, 80, 82, 83, 84, 85, 86, 88, 89, 90, 92, 93, 95, 96, 110, 113, 117, 118, 119, 120, 121, 122, 123, 125, 132, 133, 134, 151, 156, 157, 158, 159, 161, 166, 180, 181, 189, 190, 191, 196	1, 2, 10, 11, 12, 13, 14, 24, 29, 30, 41, 42, 54, 59, 73, 82, 83, 84, 85, 86, 88, 89, 90, 92, 93, 110, 113, 117, 118, 119, 120, 121, 125, 132, 133, 151, 156, 157, 161, 166, 180, 181, 189, 191	1, 10, 12, 14, 24, 29, 30, 41, 42, 54, 59, 82, 84, 85, 86, 88, 89, 90, 92, 93, 113, 117, 118, 119, 120, 121, 125, 132, 133, 156, 157, 161, 166, 189, 191
3	19, 20, 26, 53, 57, 58, 68, 69, 91, 109, 116, 129, 131, 162, 168, 176, 200	53, 57, 68, 69, 91, 109, 129	53, 57, 68, 69, 91, 109, 129
4	3, 4, 23, 25, 27, 38, 39, 56, 60, 61, 62, 114, 115, 127, 152, 155, 172, 195	3, 4, 23, 25, 27, 38, 39, 56, 60, 62, 114, 115, 127, 152, 155, 172, 195	3, 4, 23, 25, 27, 38, 39, 56, 60, 62, 114, 115, 127, 152, 155, 172, 195
5	46, 49, 100, 101, 102, 112, 177, 197	46, 49, 100, 101, 102, 112, 177, 197	46, 49, 100, 101, 102, 112, 177, 197
6	6, 7, 8, 21, 22, 51, 52, 64, 128, 130, 138, 139, 144, 160, 164, 165, 167	7, 21, 22, 64, 138, 139, 144, 160, 164, 165, 167	21, 22, 64, 138, 139, 160, 164, 167
7	70, 111, 142, 145, 171, 173, 199	70, 111, 142, 145, 171, 173, 199	70, 111, 142, 145, 171, 173, 199
8	15, 16, 94, 97, 124, 135, 178	6, 8, 15, 16, 19, 20, 26, 37, 47, 50, 51, 52, 58, 61, 66, 67, 71, 72, 75, 77, 78, 79, 80, 87, 94, 95, 96, 97, 99, 106, 116, 122, 123, 124, 128, 130, 131, 134, 135, 137, 158, 159, 162, 163, 168, 176, 178, 182, 184, 190, 196, 198, 200	19, 20, 26, 58, 116, 131, 162, 176, 200
9	5, 9, 65, 146, 153, 154, 193, 194	5, 9, 65, 146, 153, 154	2, 11, 13, 37, 73, 83, 110, 122, 123, 134, 151, 158, 180, 181, 190
10	40, 48, 79, 98, 103, 104, 108, 183	40, 48, 98, 103, 104, 108, 183	40, 48, 79, 80, 95, 96, 98, 103, 104, 108, 168, 183
11	105, 107	105, 107	5, 9, 65, 146, 153, 154, 193, 194
12	28, 43, 126	193, 194	6, 7, 8, 51, 52, 61, 128, 130, 144, 159, 165, 196
13	47, 50, 66, 67, 71, 72, 75, 77, 78, 87, 99, 106, 137, 163, 182, 184, 198	28, 43, 126	50, 63, 66, 67, 71, 72, 74, 75, 76, 77, 78, 136, 137, 163, 182, 184
14	17, 18, 81	17, 18, 81	105, 107
15	74	74	28, 43, 126
16	63	63	15, 16, 94, 97, 124, 135, 178
17	76	76	17, 18, 81
18	136	136	47, 87, 99, 106, 198
19	140	140	140
20	169	169	169

Table 2. Comparison of cluster analysis results for 1992.

Cluster	Average Linkage	Centroid Hierarchical	Ward's Minimum Variance
1	31, 33, 36, 147, 148, 175	31, 33, 36, 147, 148, 175	31, 33, 36, 147, 148, 175
2	1, 10, 11, 12, 13, 14, 24, 29, 30, 41, 42, 54, 59, 73, 82, 88, 89, 90, 92, 93, 113, 117, 118, 119, 120, 121, 123, 125, 132, 133, 156, 157, 158, 161, 166, 180, 181, 189, 191	1, 2, 10, 11, 12, 13, 14, 24, 29, 30, 41, 42, 54, 59, 73, 82, 88, 89, 90, 92, 93, 113, 117, 118, 119, 120, 121, 123, 125, 132, 133, 156, 157, 158, 161, 166, 180, 181, 189, 191	1, 10, 11, 12, 13, 14, 24, 29, 30, 41, 42, 54, 59, 73, 82, 88, 89, 90, 92, 93, 113, 117, 118, 119, 120, 121, 123, 125, 132, 133, 156, 157, 158, 161, 166, 180, 181, 189, 191
3	3, 4, 19, 20, 23, 25, 26, 27, 38, 39, 56, 57, 58, 60, 62, 68, 69, 91, 109, 114, 115, 116, 127, 129, 130, 131, 152, 162, 168, 172, 195, 200	3, 4, 23, 27, 38, 39, 60, 62, 127, 152, 172, 195 ,	3, 4, 23, 27, 38, 39, 60, 62, 127, 152, 172, 195
4	70, 77, 111, 142, 145, 151, 171, 173, 199	70, 77, 111, 142, 145, 151, 171, 173, 199	70, 77, 111, 142, 145, 151, 171, 173, 199
5	46, 49, 100, 101, 102, 110, 112, 177, 197	46, 49, 100, 101, 102, 110, 112, 177, 197	46, 49, 100, 101, 102, 110, 112, 177, 197
6	6, 7, 17, 21, 22, 35, 47, 50, 51, 61, 64, 66, 67, 71, 72, 74, 75, 78, 99, 103, 106, 128, 137, 138, 139, 140, 141, 144, 159, 160, 163, 164, 165, 167, 182, 183, 184, 196, 198	6, 7, 15, 16, 17, 19, 20, 21, 22, 25, 26, 37, 40, 48, 50, 51, 56, 57, 58, 64, 66, 67, 68, 69, 71, 72, 74, 75, 78, 79, 80, 91, 94, 95, 96, 97, 98, 104, 105, 108, 109, 114, 115, 116, 122, 124, 128, 129, 130, 131, 134, 135, 137, 138, 139, 140, 141, 144, 159, 160, 162, 163, 164, 165, 167, 168, 176, 178, 182, 183, 184, 190, 196, 200	6, 7, 21, 22, 51, 64, 128, 138, 139, 159, 160, 164, 165, 167, 184, 196
7	2, 15, 16, 37, 40, 48, 79, 80, 94, 95, 96, 97, 98, 104, 105, 108, 122, 124, 134, 135, 176, 178, 190	35, 47, 61, 99, 103, 106, 198	15, 16, 37, 94, 95, 96, 97, 122, 124, 134, 135, 176, 178, 190
8	5, 65, 146, 153	5, 65, 146, 153	26, 57, 58, 68, 69, 91, 109, 129, 162, 168, 200
9	28, 43, 126	28, 43, 126	19, 20, 25, 56, 114, 115, 116, 130, 131
10	193, 194	193, 194	2, 40, 48, 79, 80, 98, 104, 105, 108
11	63	63	5, 65, 146, 153
12	107	107	28, 43, 126
13	76	76	35, 47, 61, 99, 103, 106, 183, 198
14	18	18	193, 194
15	81	81	50, 63, 66, 67, 72, 74, 75, 76, 78, 107, 136, 137, 140, 141, 163, 182
16	136	136	17, 18, 71, 81, 144
17	8	8	8
18	154	154	154
19	169	169	169
20	52	52	52

Table 3. Vegetation types identified at White Sands Missile Range.

Types	No. of plots	Plot ID	Major species
1	3	36, 147, 175	ALOC2, PSAR, ATCA2
2	35	1, 10, 12, 14, 24, 29, 30, 41, 42, 54, 59, 82, 84, 85, 86, 88, 89, 90, 92, 93, 113, 117, 118, 119, 121, 125, 132, 133, 156, 157, 161, 166, 189, 191	ATCA2
3	7	53, 57, 68, 69, 91, 109, 129	FLCE
4	17	3, 4, 23, 25, 27, 38, 39, 56, 60, 62, 114, 115, 127, 152, 155, 172, 195	LATR2
5	8	46, 49, 100, 101, 102, 112, 177, 197	ARFI2, EPTO
6	8	21, 22, 64, 138, 139, 160, 164, 167	PAIN2, LATR2, VIST, DAFO
7	7	70, 111, 142, 145, 171, 173, 199	DAFO, PAIN2, LATR2
8	9	19, 20, 26, 58, 116, 131, 162, 176, 200	FLCE, LATR2
9	14	2, 11, 13, 37, 73, 83, 122, 123, 134, 151, 158, 180, 181, 190	ATCA2, LYBE
10	6	15, 94, 97, 124, 135, 178	LYBE, ATCA2
11	8	5, 9, 65, 146, 153, 154, 193, 194	VIST, DAFO, ACCO2, PAIN2
12	12	6, 7, 8, 51, 52, 61, 128, 130, 144, 159, 165, 196	PAIN2, LATR2, DAFO
13	12	40, 48, 79, 80, 95, 96, 98, 103, 104, 108, 168, 183	EPTO, ATCA2
14	18	50, 63, 66, 67, 71, 74, 75, 76, 77, 78, 105, 107, 136, 137, 140, 163, 182, 184	DAWH2, PAIN2, YUBA
15	3	17, 18, 81	OPVI, ACGR
16	4	47, 87, 106, 198	YUEL, POIN3
17	1	169	DAFO, PAIN2, FLCE
18	27	16, 28, 31, 33, 35, 43, 72, 99, 110, 120, 126, 150 32, 34, 44, 45, 55, 143, 148, 149, 174, 179, 185, 186, 187, 188, 192	Bare land

Simulation Results

The STELLA model as described in Chapter 5 was used to model the population and diversity of the 17 plant community types. The data used for the simulation were from four years: 1989, 1990, 1991, and 1992. Figures 4 through 20 illustrate results of the simulation for the different plant communities.

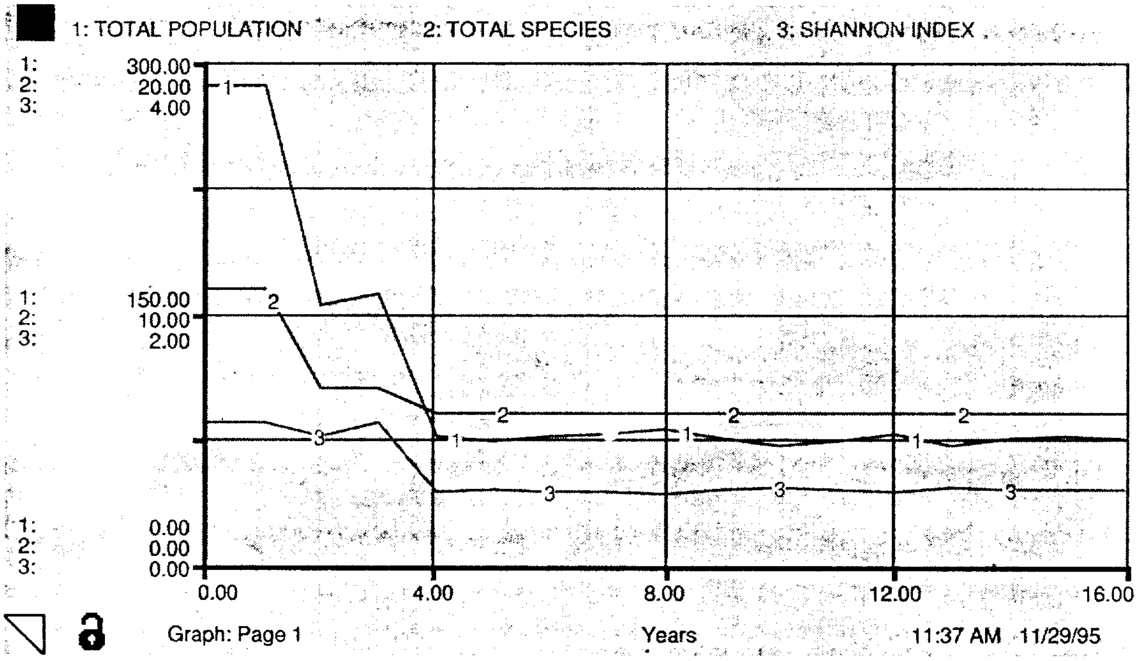


Figure 4. Simulation results for plant community type 1.

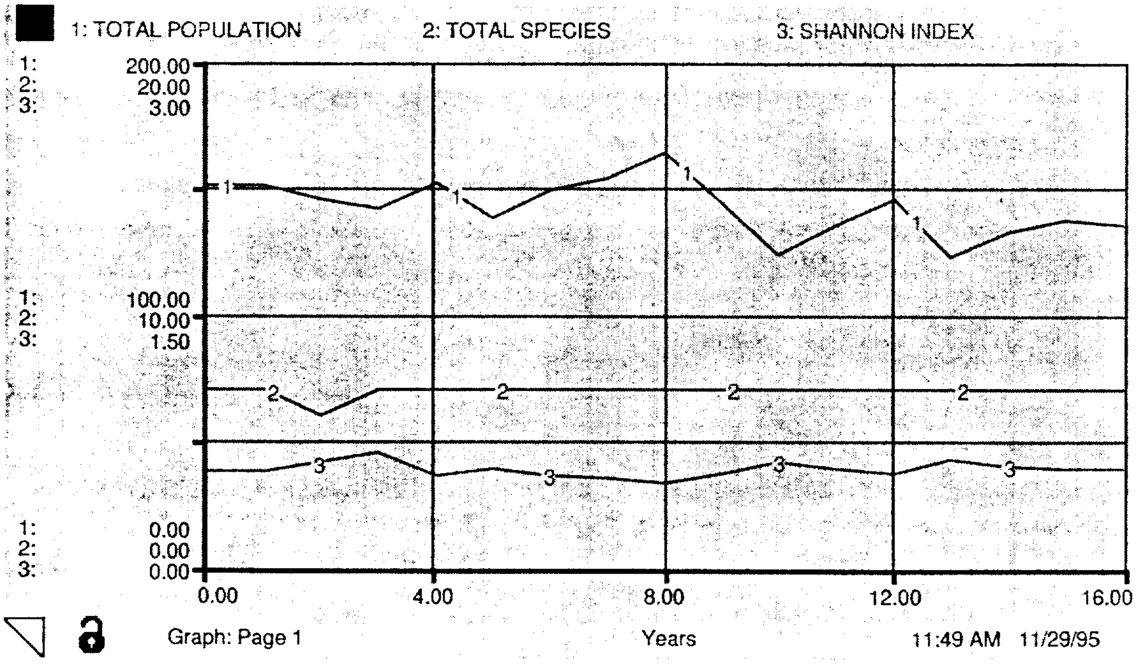


Figure 5. Simulation results for plant community type 2.

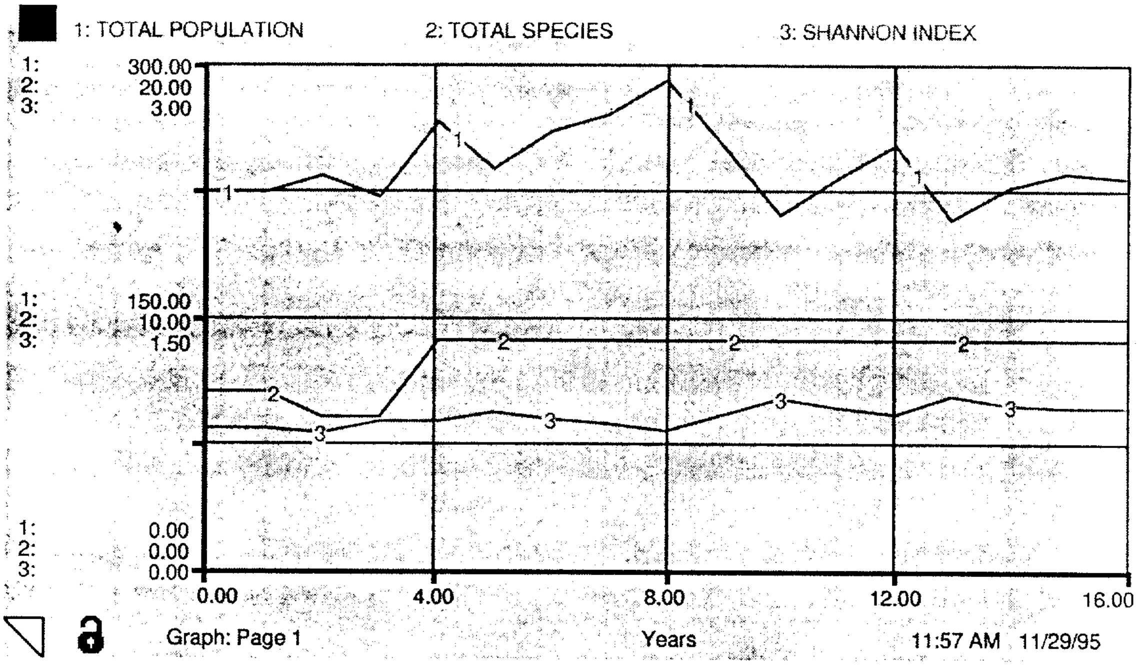


Figure 6. Simulation results for plant community type 3.

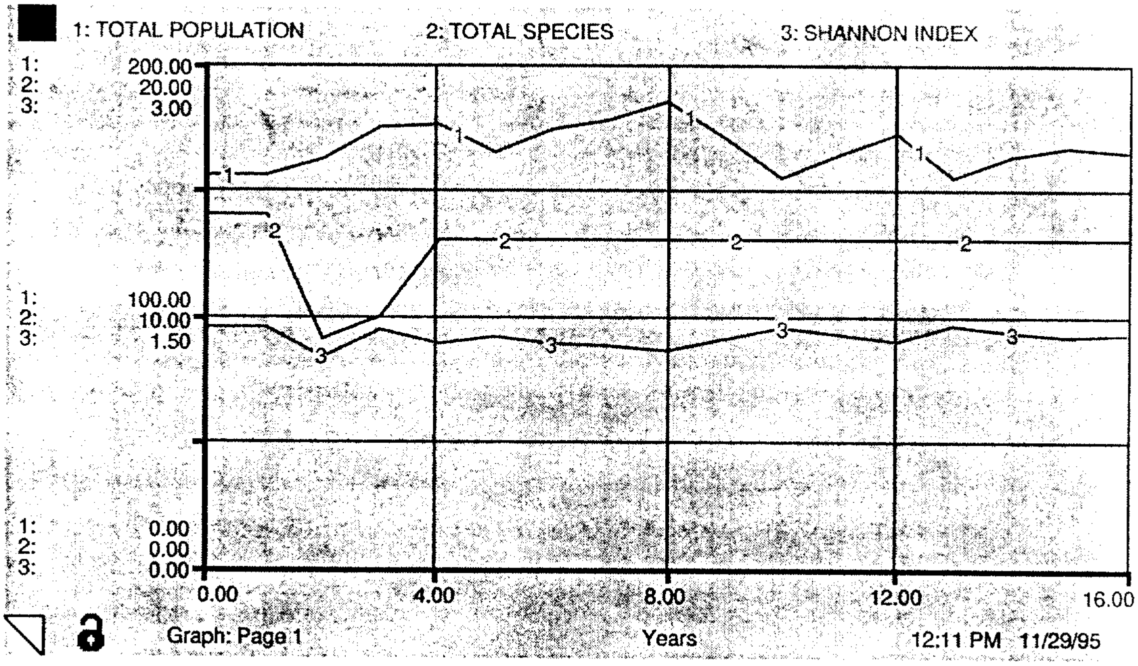


Figure 7. Simulation results for plant community type 4.

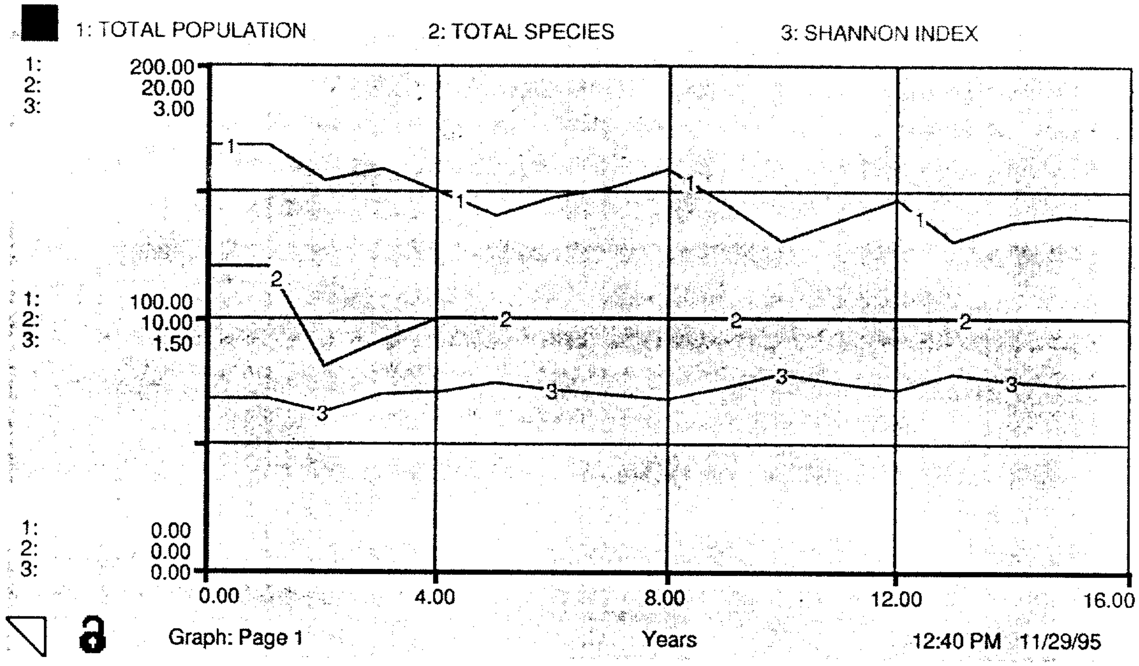


Figure 8. Simulation results for plant community type 5.

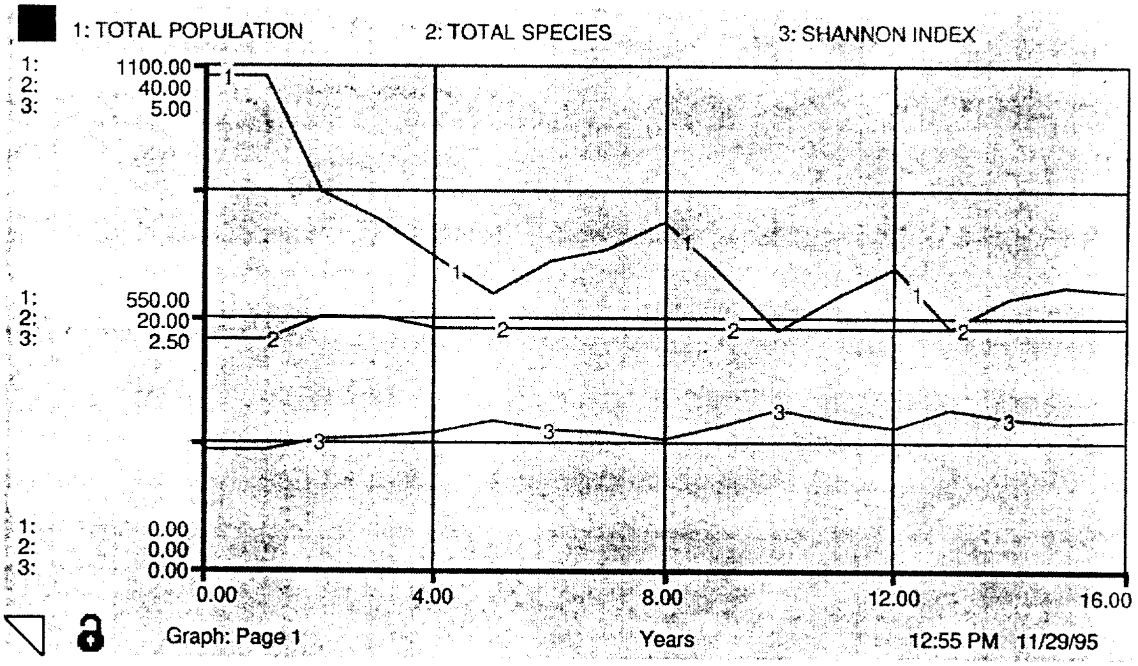


Figure 9. Simulation results for plant community type 6.

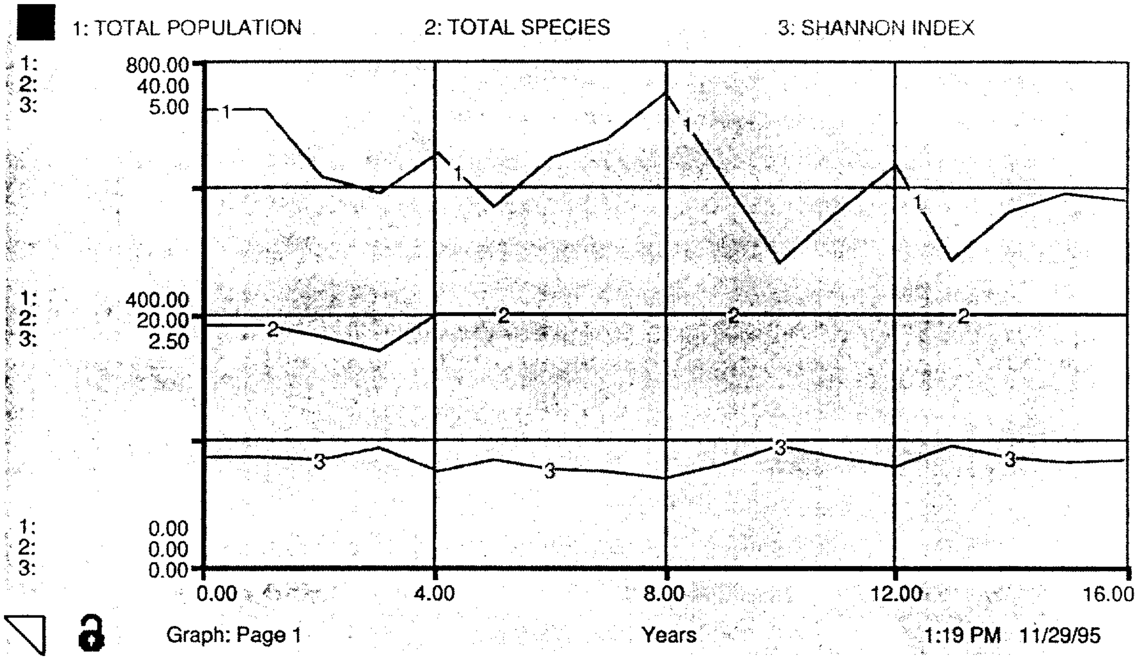


Figure 10. Simulation results for plant community type 7.

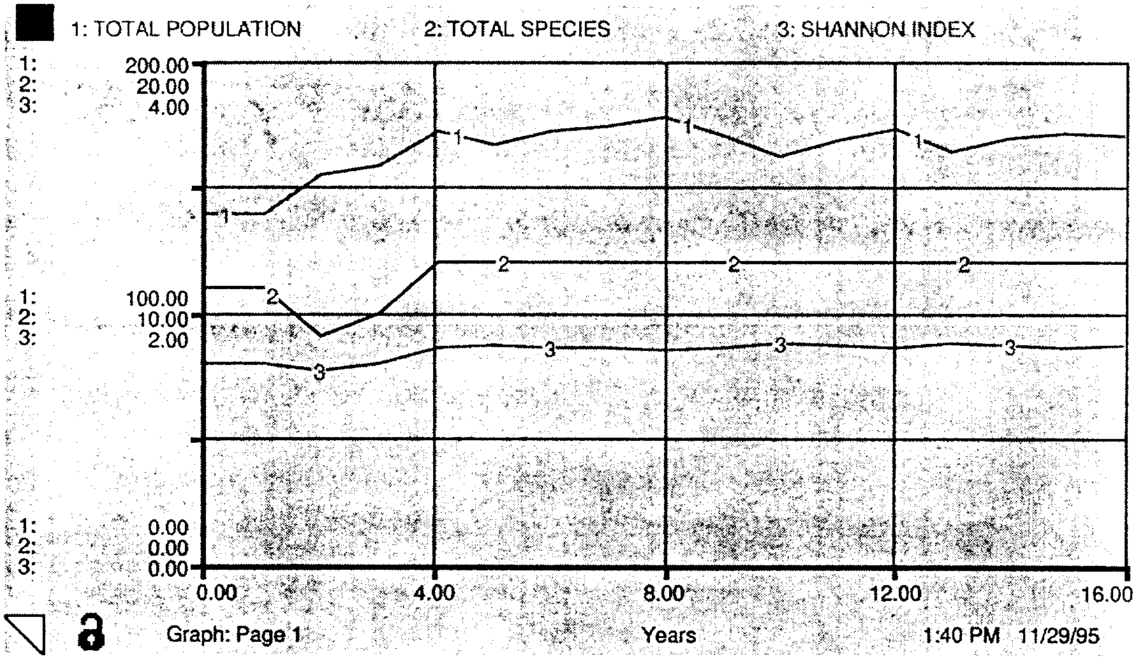


Figure 11. Simulation results for plant community type 8.

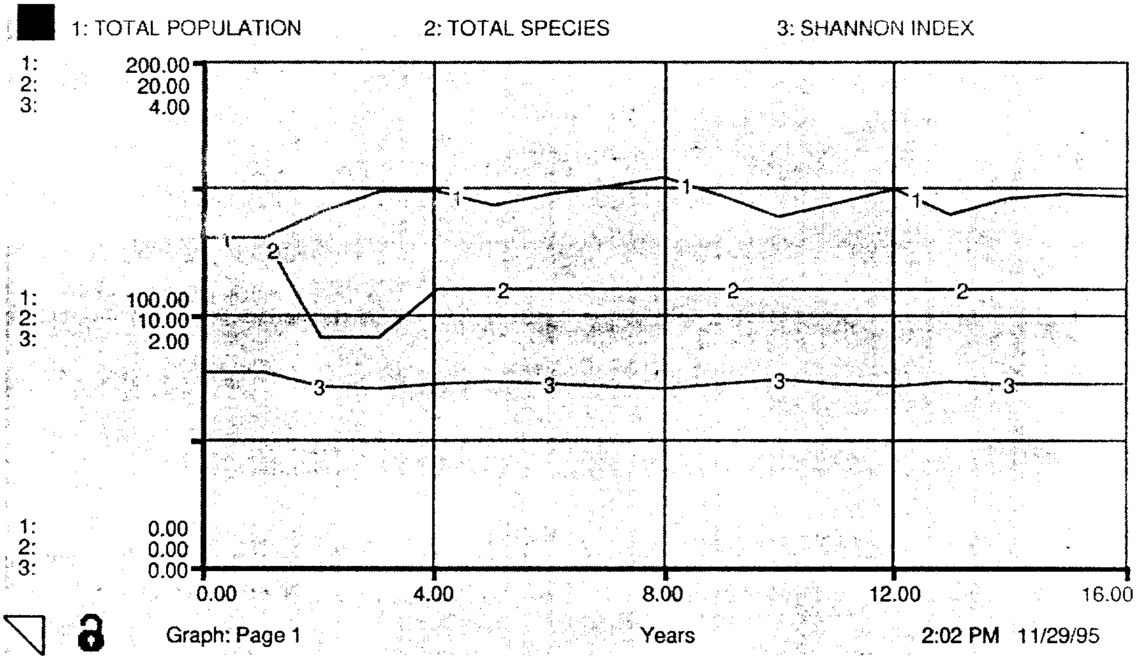


Figure 12. Simulation results for plant community type 9.

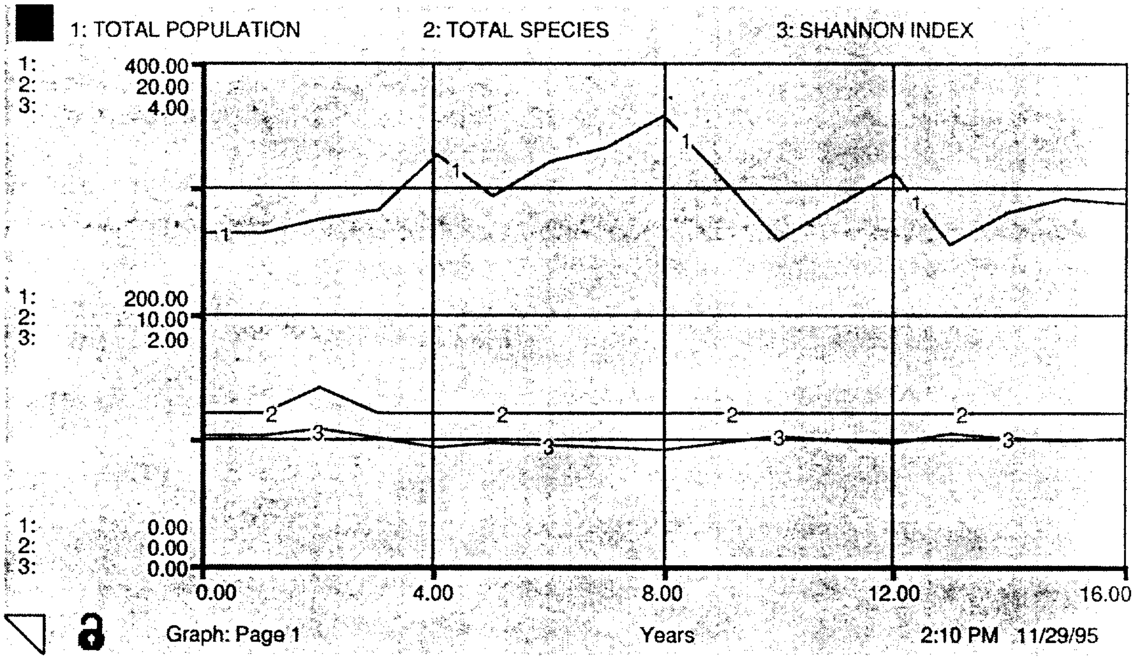


Figure 13. Simulation results for plant community type 10.

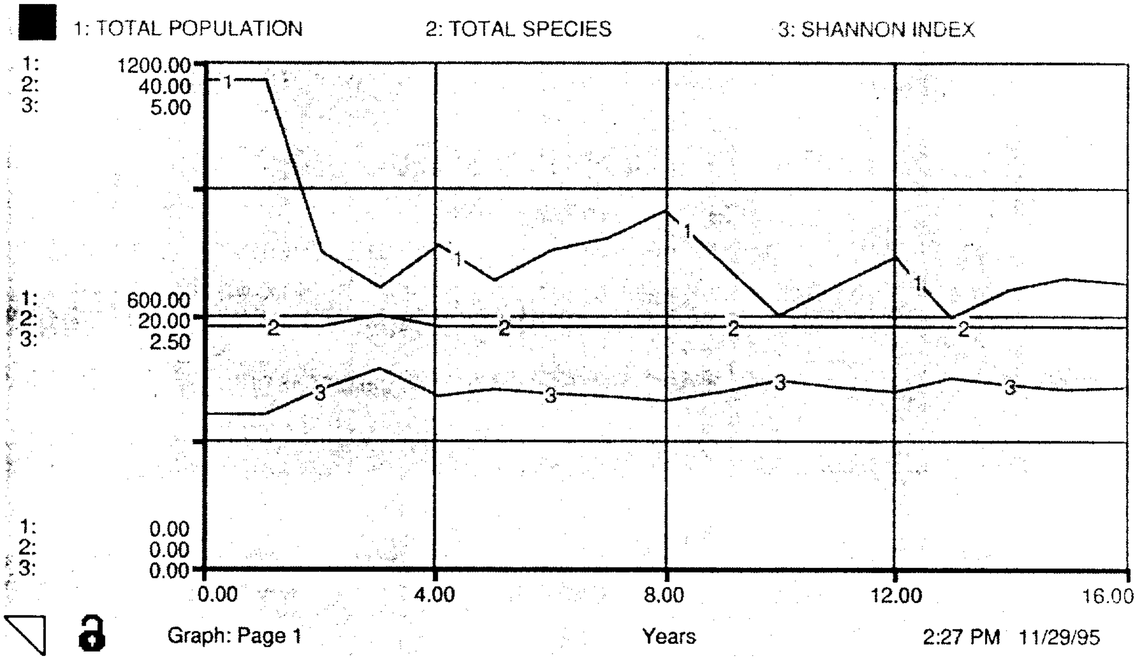


Figure 14. Simulation results for plant community type 11.

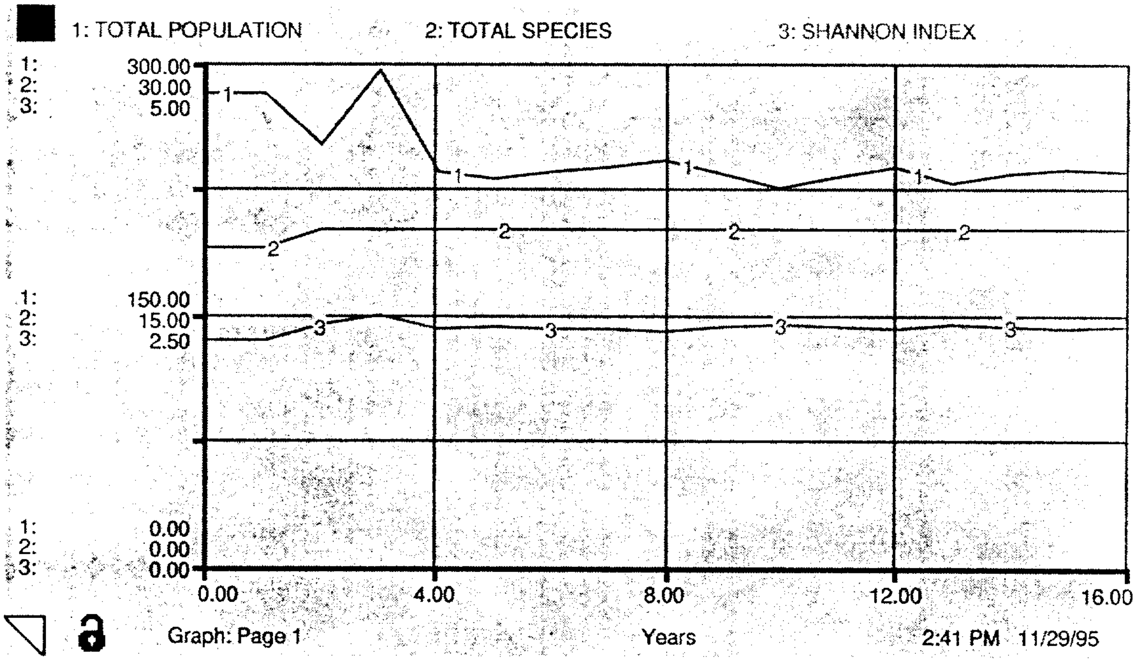


Figure 15. Simulation results for plant community type 12.

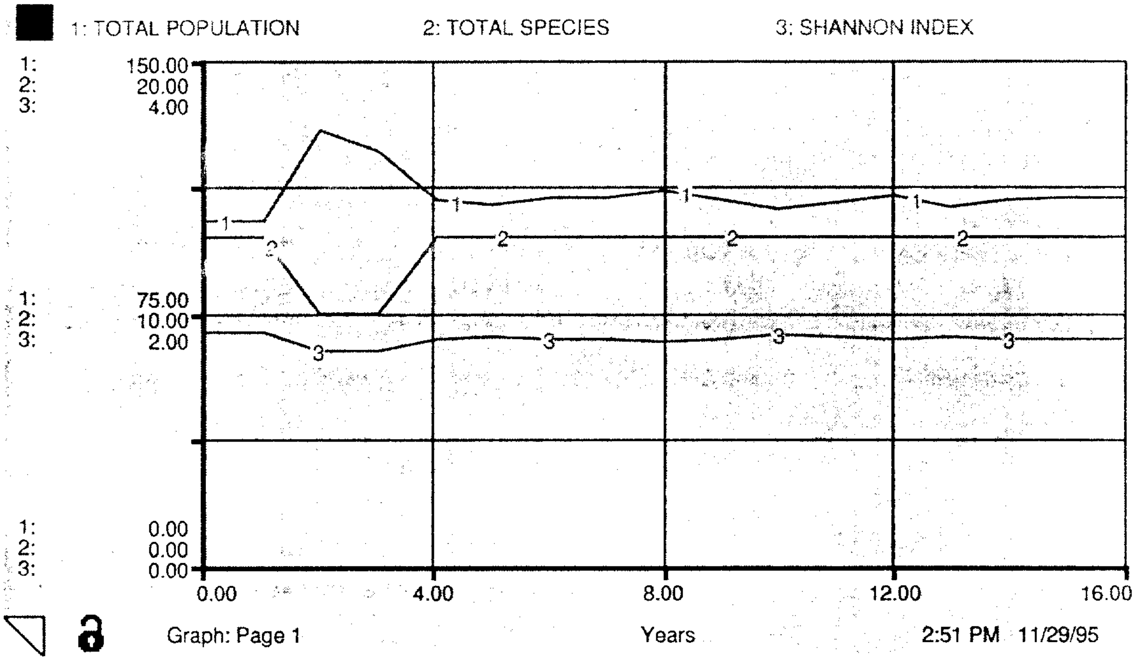


Figure 16. Simulation results for plant community type 13.

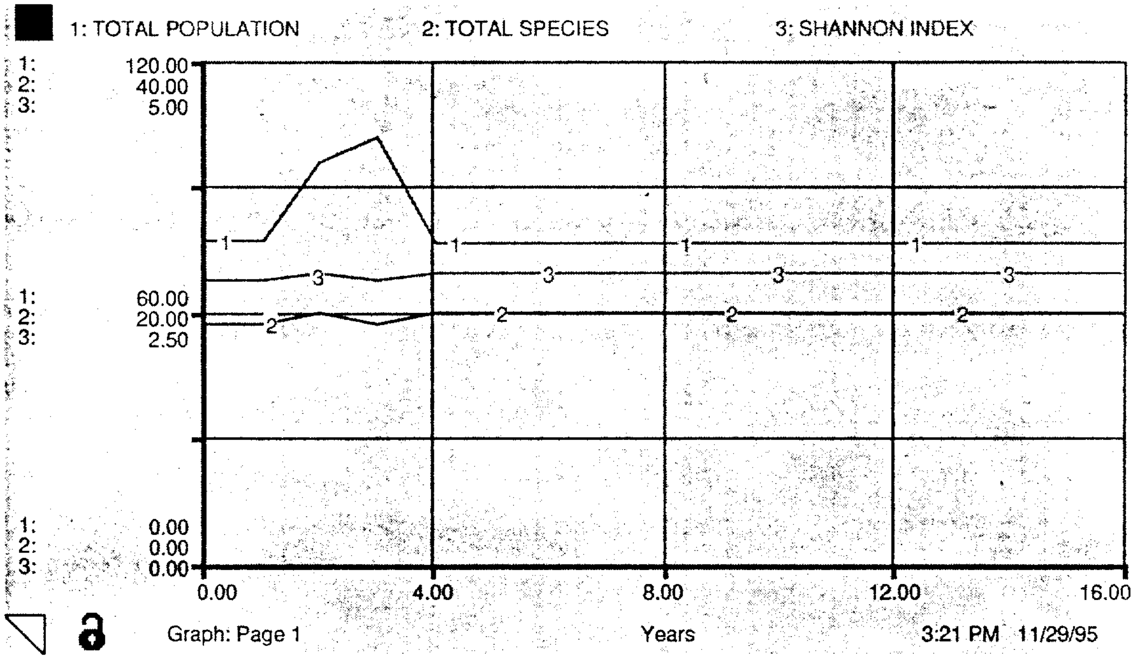


Figure 17. Simulation results for plant community type 14.

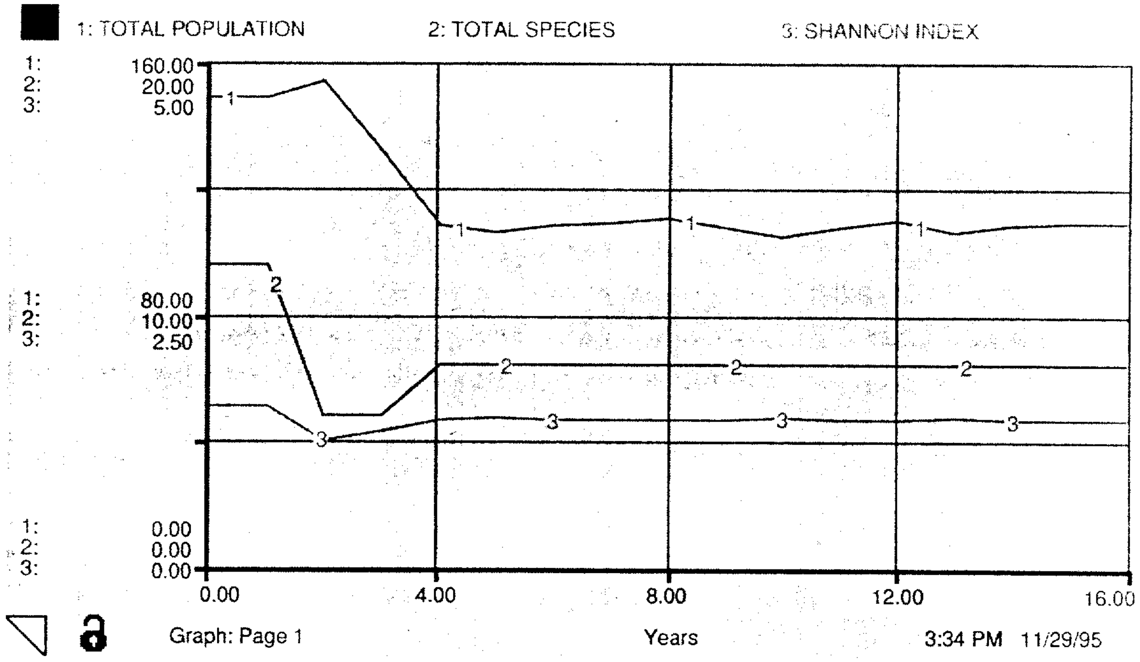


Figure 18. Simulation results for plant community type 15.

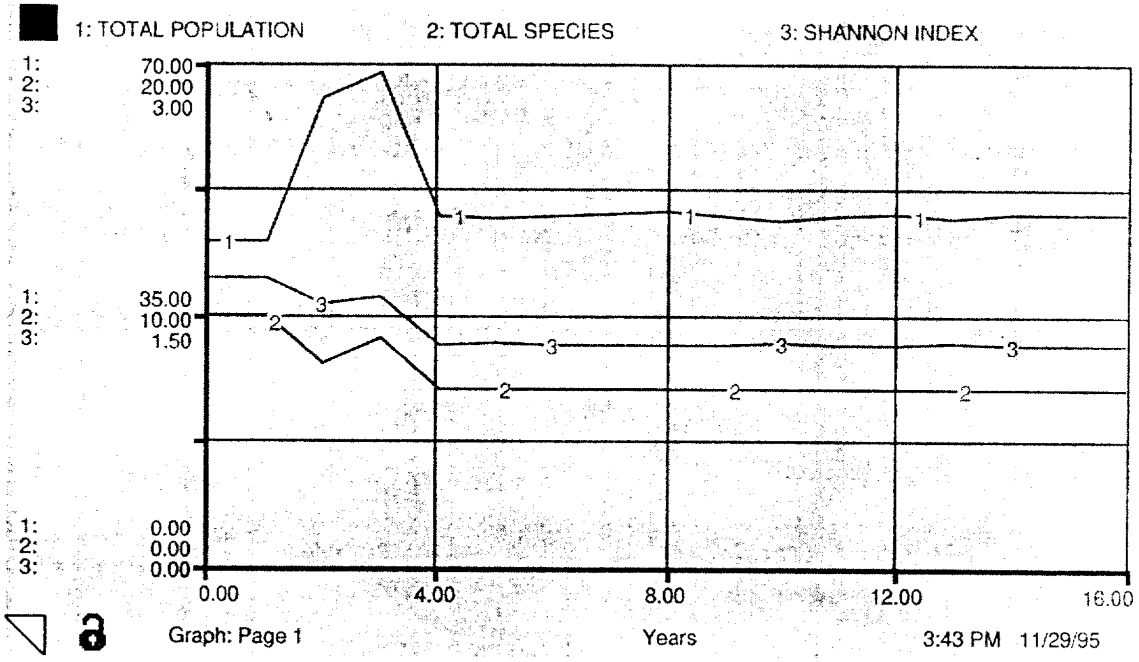


Figure 19. Simulation results for plant community type 16.

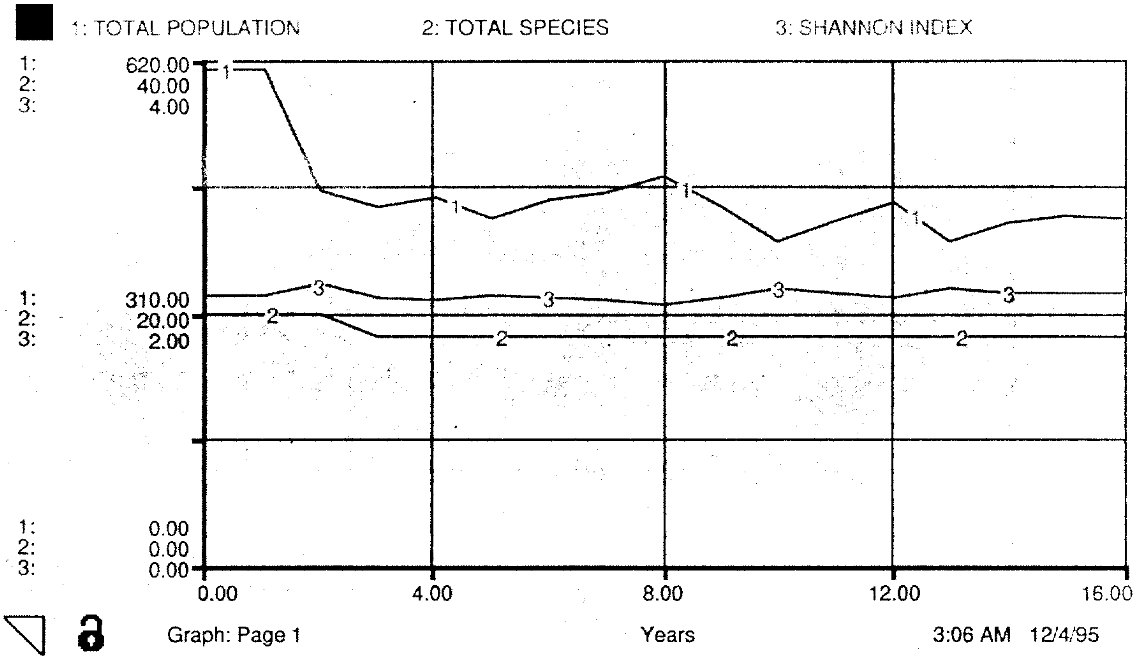


Figure 20. Simulation results for plant community type 17.

7 Summary

This report is the culmination of a project that was designed to develop and test a new methodology to model changes in plant diversity. Using standard data from the U.S. Army's LCTA program at White Sands Missile Range, New Mexico, stochastic models of plant diversity were created to simulate the dynamics of the population growth in a number of plant communities on the installation.

This project resulted in the development of a new model in which we demonstrated different methods to solve a variety of stochastic differential equations. We illustrated how to incorporate both demographic noise and environmental noise into a single model containing the joint effects of demographic and environmental stochasticity. This new model contains two main components: deterministic process and stochastic process. We first introduced a dynamics model of population growth. This model derives birth and death rates as they relate to population growth from the relationships among plants. It also derives these rates from the relationship between plants and the environment. This model serves as the deterministic part of the stochastic model and controls the biological dynamics of the plant communities. The second component of the model simulates the biological and ecological fluctuations. In the new model, the stochastic process is simulated with the birth and death process. This process best describes the demographic stochasticity because it depicts population dynamics with biologically accurate, interpretable birth and death rates and is applicable to populations of varying sizes.

During the testing phase of the project, we needed to have clearly defined plant communities in order to test the model with White Sands LCTA data. We used Ward's minimum variance method of cluster analysis because it yielded the best match of plots out of the three methods of cluster analysis used to characterize individual plant communities on the White Sands installation.

This diversity model complements other models that use LCTA data to determine plant population levels. Plant population models are useful tools in that they enable natural resource managers to transcend current identification strategies and can help determine future training levels that will allow the maximum level of training to occur in areas with minimal impact on overall species diversity. Although this particular model is only applicable at the White Sands installation, this new model can handle a variable number of species and species abundance. This capability enables it to be used to calculate any kind of diversity indices. Because both the dynamic changes and stochastic fluctuations are included in the stochastic dynamics model, re-parameterizing the model gives it the potential for extensive use in natural resource management and environmental monitoring. Future applications of these stochastic dynamics models include: providing standard diversity measures; monitoring the development of plant communities in terms of species diversity and structure diversity; testing the significance of the influence of human activities on plant communities; and estimating rehabilitation time for a disturbed plant community, thereby helping the Army integrate training and natural resource management.

References

Bartlett, M.S. 1966. *An Introduction to Stochastic Processes*. Cambridge University Press.

Bechtel, T.J., and B.J. Copeland. 1970. "Fish species diversity indices as indicators of pollution in Galveston Bay, Texas." *Contrib. Mar. Sci.*, 15, 103-32.

Berger, W.H., and F.L. Parker. 1970. "Diversity of planktonic Foraminifera in deep sea sediments." *Science*, 168, 1345-7.

Bharucha-Reid, A.T. 1960. *Elements of the Theory of Markov Processes and Their Applications*. McGraw-Hill, New York.

Capocelli, R.M., and L.M. Ricciardi. 1974. "A diffusion model for population growth in a random environment." *Theoretical Population Biology*. 5, 28-41.

Caswell, H. 1976. "Community structure: a neutral model analysis." *Ecol. Monogr.*, 46, 327-354.

Clifford, H.T., and W. Stephenson. 1975. *An Introduction to Numerical Classification*. Academic Press, London.

Dennis, B., and G.P. Patil. 1979. "Species abundance, diversity, and environmental predictability." In *Ecological Diversity in Theory and Practice*, J.F. Grassle, G.P. Patil, W.K. Smith, and C. Taillie, eds. International Cooperative Publishing House, Fairland, Maryland.

Diersing, V.E., R.B. Shaw, and D.J. Tazik. 1992. "U.S. Army Land Condition Trend Analysis (LCTA) Program." *Environmental Management,* 16:405-414.

Elton, C.S., and R.S. Liller. 1954. "The ecological survey of animal communities: with practical system of classifying habitats by structural characteristics." *J. Ecol.*, 42, 460-96.

Engen, S. 1978. *Stochastic Abundance Models with Emphasis on Biological Communities and Species Diversity*. Chapman and Hall, London.

Fisher, R.A., A.S. Corbet, and C.B. Williams. 1943. "The relationship between the number of species and the number of individuals in a random sample of an animal population." *J. Anim. Ecol.*, 12, 42-58.

Fuller, R.J., and D.R. Langslow. 1986. "Ornithological evaluation for wildlife conservation." In *Wildlife Conservation Evaluation*, M.B. Usher, ed., Chapman and Hall, London.

Gelb, A. 1974. *Applied Optimal Estimation*. The M.I.T. Press, Cambridge, Massachusetts.

Gertner, G.Z., X. Cao, and H. Zhu. 1995. "A Quality Assessment of A Weibull Based Growth Projection System," *Forest Ecology and Management.*

Goel, N.S., and N. Richer-Dyn. 1974. *Stochastic Models in Biology.* Academic Press, New York.

Gray, J.S. 1988. "Species abundance patterns." In *Organization of Communities: Past and Present.* Blackwell, Oxford.

Gray, J.S. 1978. "The structure of meiofauna communities." *Sarsia*, 64, 265-272.

Gray, J.S., and F.B. Mirza. 1979. "A possible method for detecting pollution-induced disturbance on marine benthic communities." *Mai. Pollut. Bull.*, 10, 142-6.

Hallam, T. 1986. "Population Dynamics in a Homogeneous Environment." *Biomathematics, Volume 17, Mathematical Ecology,* T.G. Hallam and S.A. Levin, eds. Springer, Berlin.

Homer, M. 1976. "Seasonal abundance, biomass, diversity and trophic structure of fish in a saltmarsh tidal creek affected by a coastal power plant." In *Thermal Ecology II*, G.W. Esch and R.W. McFarlance, eds. U.S. Energy Research and Development Administration, Washington, DC.

Hughes, R.G. 1984. "A model of the structure and dynamics of benthic marine invertebrate communities." *Mar. Ecol. Prog. Ser.*, 15, 1-11.

Hughes, R.G. 1986. "Theories and models of species abundance." *Amer. Nat.*, 128, 879-99.

Hurlbert, S.H. 1971. "The non-concept of species diversity: a critique and alternative parameters." *Ecology,* 52, 577-86.

Judge, G.G., [et al], 1988. Introduction to the theory and practice of econometrics. John Wiley & Sons, Inc., New York.

Kalman, R.E. 1960. "A new approach to linear filtering and prediction problems." *Transaction of the ASME, Series D,* 82, 35-45.

Kempton, R.A., and R. Wedderburn. 1979. "A comparison of three measures of species diversity." *Biometrics*, 34, 25-37.

Kempton, R.A., and L.R. Taylor. 1974. "Log-series and log-normal parameters as diversity discriminants for the Lepidoptera." *J. Anim. Ecol.*, 43, 381-399.

Kershaw, K.A., and J. Looney. 1985. *Quantitative and Dynamics Plant Ecology.* Arnold, London.

Magurran, A.E. 1988. *Ecological Diversity and Its Measurement.* Princeton University, Princeton, New Jersey.

Mandelbrot, B.B. 1977. *Fractals, Fun, Chance and Dimension.* W.H. Freeman, San Francisco.

May, R.M. 1973. *Stability and Complexity in Model Ecosystems.* Princeton University Press, Princeton, New Jersey.

May, R.M. 1974. "General introduction." In *Ecological Stability*, M.B. Usher and M.H. Williamson, eds, Chapman and Hall, London.

May, R.M. 1975. "Patterns of species abundance and diversity." In *Ecology and Evolution of Communities*, M.L. Cody and J.M. Diamond, eds, Harvard University Press, Cambridge, MA.

MacArthur, R.H. 1957. "On the relative abundance of bird species." *Proc. Nat. Acad. Sci.*, USA, 43, 293-5.

MacArthur, R.H., and J.W. MacArthur. 1961. "On bird species diversity." *Ecology*, 42, 594-8.

McIntosh, R.P. 1967. "An Index of Diversity and the Relation to Certain Concepts of Diversity." *Ecology,* 48, 392-404.

Nisbet, R.M., and W. Gurney. 1982. *Modeling Fluctuating Populations*. John Wiley and Sons, New York.

Otter, P.W. 1978. "The discrete Kalman filter applied to linear regression models: Statistical considerations and an application." *Statistica Neerlandica* 32, 41-56.

Pearsall, S.H., D. Durham, and D.C. Eagar. 1986. "Evaluation methods in the United States." In *Wildlife Conservation Evaluation*, M.B. Usher, ed., Chapman and Hall, London.

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

Preston, F.W. 1948. "The commonness and rarity of species." *Ecology,* 29, 254-283.

Preston, F.W. 1962. "The canonical distribution of commonness and rarity." *Ecology*, 43, 185-215, 410-432.

Preston, F.W. 1980. "Noncanonical distribution of commonness and rarity." *Ecology*, 61, 88-97.

Rose, C.I. 1978. "A note on diversity and conservation." *Bull. Brit. Ecol. Soc.*, 9(4), 5-6.

Ross, S.M. 1983. *Stochastic Process*. Wiley and Sons, New York.

Rozovskii, B.L. 1990. *Stochastic Evolution Systems*. Kluwer Academic Publishers, Norwell, MA.

Sander, H.L. 1968. "Marine benthic diversity: a comparative study." *Amer. Nat.*, 102, 243-82.

Schafer, C.T. 1973. "Distribution of foraminifera near pollution sources in Chaleur Bay." *Water Air Soil Pollution*, 2, 219-33.

Shaw, R.B., and D.G. Kowalski. 1996. *U.S. Army Lands: A National Survey*, The Center for Ecological Management of Military Lands CEMML TPS 96-1 (Colorado State University, Fort Collins, CO).

Simpson, E.H. 1949. "Measurement of diversity." *Nature,* 163, 688.

Southwood, T., V.K. Brown, and P.M. Reader. 1979. "The relationship of plant and insect diversities in succession." *Bio. J. Linn. Soc.*, 12, 327-48.

STELLA II, Technical Documentation (High Performance Systems, Inc., Hanover, NH, 1994).

Streifer, W. 1974. "Realistic Models in Population Ecology." *Advances in Ecological Research*, 8, 199-266.

Taylor, L.R., R.A. Kempton, and I.P. Woiwod. 1976. "Diversity statistics and the log-series model." *J. Anim. Ecol.*, 45, 255-272.

Technical Note [TN] 420-74-3. 1990. *Army Land Inventory and Monitoring Procedures on Military Installations* (U.S. Army Engineering and Housing Support Center [USAEHSC], Fort Belvoir, VA).

Tomascik, T., and F. Sander. 1987. "Effects of eutrophication on reef building corals." *Mar. Biol.*, 94, 53-75.

Tuckwell, H.C. 1974. "A study of some diffusion models of population growth." *Theoretical Population Biology*, 5, 345-357.

Turelli, M. 1977. "Random environment and stochastic calculus." *Theoretical Population Biology*, 12, 140-178.

Turelli, M. 1986. "Stochastic community theory: a partially guided tour." In *Mathematical Ecology,* T.G. Hallam and S.A. Levin, eds, Springer-Verlag, London.

Usher, M.B. 1986. "Wildlife conservation evaluation: attributes, criteria and values." In *Wildlife Conservation Evaluation*, M.B. Usher, ed., Chapman and Hall, London.

Visser, H., and J. Molenaar. 1990. "Estimating Trends in Tree-ring Data." *For. Sci.* 36(1): 87-100.

Whittaker, R.H. 1965. "Dominance and diversity in land plant communities." *Science*, 147, 250-60.

Whittaker, R.H. 1970. *Communities and Ecosystems*. Macmillan, New York.

Whittaker, R.H. 1972. "Evolution and measurement of species diversity." *Taxon*, 21, 213-51.

Williams, C.B. 1964. "Patterns in the Balance of Nature and Related Problems." In *Quantitative Ecology,* Academic Press, London.

Wu, R. 1982. "Periodic defaunanation and recovery in a subtropical epibenthic community, in relation to organic pollution." *J. Exp. Mar. Biol. Ecol.*, 64, 253-69.

Yapp, W.B. 1979. "Specific diversity in woodland birds." *Field Stud.*, 64, 253-69.

Zipf, G.K. 1965. *Human Behavior and the Principle of Least Effort.* Hafner, New York.

DISTRIBUTION

Chief of Engineers
    ATTN: CEHEC-IM-LH (2)
    ATTN: HECSA Mailroom (2)
    ATTN: CECC-R
    ATTN: CERD-L
    ATTN: CERD-M

Defense Tech Info Center 22304
    ATTN: DTIC-O (2)

ACS (IM)
    ATTN: DAIM-ED-N (2)

11
10/99

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave Blank*)	2. REPORT DATE February 2000	3. REPORT TYPE AND DATES COVERED Final

4. TITLE AND SUBTITLE
Stochastic Models of Plant Diversity: Application to White Sands Missile Range

5. FUNDING NUMBERS
622720
A896
CN-TO9

6. AUTHOR(S)
Xiangchi Cao, George Z. Gertner, Bruce A. MacAllister, and Alan B. Anderson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Construction Engineering Research Laboratory (CERL)
P.O. Box 9005
Champaign, IL 61826-9005

8. PEFORMING ORGANIZATION REPORT NUMBER
TR 00-5

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Headquarters, Department of the Army
ATTN: DAIM-ED-N
ACS(IM)
600 Army Pentagon
Washington, DC 20310-0600

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

9. SUPPLEMENTARY NOTES

Copies are available from the National Technical Information Service, 5385 Port Royal Road, Springfield, VA 22161

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Diversity is an average property of a community and is identified as the species variety and abundance. The study of biodiversity is important because it is one of the central themes in ecology; the diversity of a system is often seen as an indicator of the well-being of the system. In this study, we selected and defined theoretical statistical measures of plant diversity and developed theoretical dynamics models for plant communities. This project provides a new approach to measuring plant diversity through time.

The stochastic dynamics modeled in this project contain two main components: deterministic processes and stochastic processes. Stochastic dynamics models make it possible to simulate plant diversity changes through time even without long-term observed data. Because both the time and space dimensions are included in these stochastic dynamics models, they have more extensive uses in assessing and monitoring plant communities. The potential applications of these models include (1) providing standard diversity measures, (2) monitoring the development of plant communities in terms of species diversity and structure diversity, (3) testing the significance of the influence of human activities on plant communities, and (4) estimating the rehabilitation time of a disturbed plant community.

14. SUBJECT TERMS

White Sands Missile Range	biodiversity	plant diversity
plant communities	LCTA	

15. NUMBER OF PAGES
87

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102