

**UNIVERSITY OF NAIROBI**

**A STATISTICAL MODEL FOR AN AQUATIC STOCK  
ESTIMATION WITH APPLICATION TO A FISH  
POPULATION**

**A RESEARCH PROPOSAL FOR THE DEGREE OF DOCTOR OF  
PHILOSOPHY IN MATHEMATICAL STATISTICS**

By

Julius Kipyegon Kones  
Department of Mathematics  
University of Nairobi  
P.O. Box 30197  
Nairobi



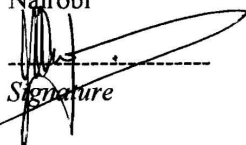
26-05-23

-----  
*Signature*

-----  
*Date*

**Supervisors**

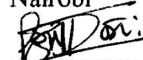
Prof. John Okoth Owino  
Department of Mathematics  
University of Nairobi  
P.O. Box 30197  
Nairobi

  
-----  
*Signature*

29-05-03

-----  
*Date*

for  
Prof. John Odhiambo  
Department of Mathematics  
University of Nairobi  
P.O. Box 30197  
Nairobi



-----  
*Signature*

27-05-03

-----  
*Date*

## **TABLE OF CONTENTS**

	<b>Page No.</b>
Title	i
Table of Contents	ii
Abstract	iii
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Background	1
1.2 Literature Review	2
1.3 Statement of the Problem	5
1.4 Aims and Objectives	6
1.5 Significance of the Study	7
<b>2 METHODOLOGY</b>	<b>7</b>
2.1 The Basic Matrix Modeling Framework	7
2.2 Data Collection and Analysis	8
2.3 Demographic Sensitivity, Elasticity and Fitness Analysis	9
<b>3 MATRIX POPULATION MODELS</b>	<b>10</b>
3.1 Leslie Matrix Approach	10
<b>4 STOCK ESTIMATION</b>	<b>11</b>
<b>5 MODELS FOR VARIABLE FISH RECRUITMENT</b>	<b>13</b>
5.1 Ricker Spawner-Recruit Model	13
5.2 Beverton and Holt Spawner-Recruit Model	14
<b>6 HARVESTING AND STOCK RECRUIT MODELS</b>	<b>15</b>
<b>7 SURPLUS PRODUCTION MODELS</b>	<b>16</b>
7.1 The Graham-Schaefer Production Model	16
7.2 Pella and Tomlison's Generalized Surplus Production Model	17
<b>8 MARKOV CHAIN MONTE CARLO METHODS</b>	<b>18</b>
8.1 Ordinary Monte Carlo Computations	18
8.2 Bayesian Computation	20
8.3 Markov Chains and Monte Carlo Calculations	20
<b>9 LIST OF REFERENCES</b>	<b>22</b>
<b>10 TIME PLAN</b>	<b>26</b>
<b>11 BUDGET</b>	<b>27</b>

## ABSTRACT

Fish stocks are renewable resources if correctly handled. If all mature fish are caught before spawning, then there will be no recruitment. The problems that relate to fish management vary enormously depending on the objective. The basic purpose of fish stock assessment is to provide advice on the optimum exploitation level of aquatic living resources such as fish. The problem is to find the appropriate middle ground where it is possible to obtain good catches for a long time. This can only be achieved through estimation of the size and productivity of fish stocks, yet it is rarely possible to count the abundance of an aquatic species. Mathematical models are used to relate the measurements to the stock size while statistical techniques are used to estimate the unknown parameters.

The study aims at developing an ecologically sustainable statistical model that can be applied to describe the dynamics of aquatic population through their life stages. The developed model will be applied to tilapia fish population in a tropical environment, with specific application to Kenyan waters. More specifically, the study will use empirical data to describe the population dynamics of tilapia fish and model it appropriately, estimate the parameters of the derived model and make recommendations that can be adopted for proper planning of fish breeding and harvesting.

The theory of matrix modeling forms the basis of the study. The Markov Chain Monte Carlo (MCMC) technique and Bayesian approach shall be applied to derive the parameters of the modeled population. Bayesian estimation makes use of prior information regarding the population under study, a property that has been lauded as the only coherent statistical methodology for updating knowledge using the information contained in the data. This property enables the posterior from one analysis to be used as an induced prior in a subsequent analysis, thereby building and exploiting an accumulated base of knowledge. The study will make use of primary and secondary data to be obtained from fisheries research institutions locally and abroad for comparative purposes. Elasticity analysis shall be incorporated in the model to indicate which parameters and ages/sizes are contributing the most to population growth rate, and which parameters have the most influence on growth rate.

# 1 INTRODUCTION

## 1.1 Background

Ecological communities are made up of a vast profusion of living things from trees to micro-organisms. Each species differs from every other. Furthermore, individuals within any one species are unique. Several factors that are known to affect the behaviour of each individual include genetic constitution, age, unpredictable changes of life up to that moment, and the prevailing conditions in the locality at the moment. There are many other influences that affect the individuals' physiological processes. A community as a whole is at all times being depleted by deaths and replenished by births. Often immigration and emigration are other causes of continual turnover of individuals. As a result the components of a community are never the same on two successive occasions. Past experience has, however, shown that in the absence of outside disturbance most communities either remain in a steady state for long periods or pass through an orderly progression of successive stages that culminates in a steady state. Two areas that are of interests to ecologists mainly are in two fold:- determining the processes that permit maintenance of a steady state, or the gradual, orderly succession of states that occur in healthy undisturbed communities, and establishing the causes and consequences of sudden departures from steadiness.

Habitat studies mainly relate to comparison of population density, reproductive rate, measures of body condition, or survival rates (Morrison et al. 1992). But measures derived from population models provide flexible, powerful and intuitively appealing methods with links to fundamental theory in population ecology and genetics (Caswell, 1989). The use of explanatory models or systems models whose behaviour is thought to duplicate (at least approximately) the true behaviour of the populations being modeled has been applied in habitat studies. This approach has its weakness in that the conclusions derived from them, though interesting and thought provoking for ecologists, only relate to simple mathematical systems and not to the complex systems, (Oster, 1975) and are not capable of discriminating among the competing theories on ecosystem's functioning. The other approach to the study of ecological systems is quantitative descriptive technique which involve manipulation

of large bodies of observational data in such a manner that will reveal interesting regularities that hitherto would be buried from sight. The revealed regularities can trigger the building of hypotheses that could as well be tested by appeal of other bodies of data. Lastly, the statistical approach is a method based on concrete, observable things, derived from real ecosystems which initially furnish an unstructured scanty of impressions that can then be used to single out those observable phenomena that serve as insights to the underlying processes. The statistical approach can therefore be said to be detective and entails discrimination of all the facts that are not feasible between the facts that are relevant and facts that are not relevant to the solution of particular problems.

## 1.2 Literature Review

Due to the profound differences in ecological systems, mathematical modeling has been found to play a key role. <sup>(in what?)</sup> In an effort to determine the processes that permits maintenance of a steady state or the gradual orderly succession of states and the causes and consequences of sudden departures from steadiness, ecologists have found the need to employ mathematical models. Population modeling is a relatively old field just as the study of ecology is (Malthus, 1798; Gompertz, 1825). The first fundamental fact of population dynamics, the idea of exponential growth of most organisms, was known as early as the 16th century (Cole, 1958). Some of the earlier works on population modeling include Lotka (1924) who developed a mathematical model of the USA population. Later Volterra (1926) independently applied mathematical models in problems of variations and fluctuations in the number of individuals and species of fish in Adriatic Sea. The results of these two works form the backbone of the modern deterministic models.

Most of the earlier models were deterministic non-linear representation of the population processes. Hutchinson (1978) pointed out that the logistic equation of limited growth, the second most fundamental aspect of population dynamics, was first derived by Verhulst (1838). Verhulst (1838) introduced the concept of limited population growth which have popularly been referred to as the logistic law (Kings-

land, 1985). The logistic equation consider all individuals as identical ignoring the demographic differences that may exist. The model by Lotka (1939) related the structure and dynamics of a homogeneous population, but this was later modified by Gause (1934) and Koztitzin (1939) to include predator-prey interaction in a system. This modification however failed to consider the effects of intraspecific competition on the wider interspecific competition and assume that time scale of the interaction amongst predator-prey population on one hand and the predation process of individual predators on the other hand is constant. Kolmogorov (1936) demonstrated the use of ordinary differential equations on single species interaction, but was also improved later by Rescigno and Richardson (1965), May (1972), and Albrecht et al. (1974) to include numerous species and interaction factors.

Deterministic matrix population models as pioneered by Bernardelli (1941), Lewis (1942) and Leslie (1945, 1948) have also been used in ecological modeling. Their stochastic versions were introduced by Bartlett (1960). Matrix populations have had applications in population modeling as in Lefkovich (1965) who introduced stages of development of a species in the models. Usher (1966, 1969) modified the model and applied it in the management of forest (and generally renewable) resources. Matrix models have also been applied in harvesting problems as in Beddington (1973) who applied it on formulation of optimal harvesting policies while Doubleday (1975) applied the model to fisheries harvesting. Matrix models can be expressed as stage (or size)-structured or as age-structured. Sauer and Slade (1987) found out that the stage structured models are more precise than the age-based in capturing the important demographic characteristics and providing insights into certain dynamical properties of the population that a Leslie matrix cannot reveal. Caswell (1989) integrated a population's dynamics into the matrix model and found out that the model is more suitable when the life cycle is described in terms of stage or size classes of development rather than the age classes which restricts the robustness of the models. On the other hand stage-based models result in potentially large numbers of parameters to be estimated (Burgman et al., 1993).

Stochastic models have also been used by various modelers to describe complex dynamics in ecological systems. Turchin (1995) used stochastic models to compare

patterns of density dependence and relative contribution of intrinsic versus extrinsic sources of variability to population dynamics. May and Oster (1976) and Hastings et al. (1993) showed that if two populations are started at initial densities the smallest fraction apart, and if the sequence of stochastic events affecting the two populations is the same, then the two populations will diverge through time if the dynamics are chaotic. But May (1973) shows that deterministic models with unstable equilibrium points show relatively low variances in population density than the stochastic models.

Fisheries modeling have tended to develop independently from general population modeling (Kerr, 1982; Werner, 1982). Fish have a highly plastic growth and span a wide range of sizes over a single age and or stage. Pauly (1980) have shown that age estimation in fish is relatively easier than other aspects of fish. This is the reason, therefore, why age-structured models of fish are sometimes preferred to stage structured ones. But on the other hand, Strong (1984) showed that size dependent interactions provide a unifying framework for understanding mechanisms governing survival and recruitment in fishes.

The Leslie (1945, 1948) age-structured model only proved useful to species with distinct ages like mammals and birds, but such models have been shown to demonstrate less robustness when applied to species like fish and insects that produce large number of offsprings and experience low survival rates during their early life history. Beddington (1984) looked into the effects of fish recruitment on harvesting and stock management. Smith (1981) noted that it is extremely difficult to precisely estimate fish mortality in the early life history stages while Gillian and Fraser (1987) developed elaborate models of fish habitat choice involving trade-offs of predation risk and growth in various habitats. Beyer (1989) found out that fish survival and fecundity depends more directly on size than any other structure due to demographic and environmental variations. A multi-structured fishery model comprising of age and stage (or size) of population species would, therefore, be more appropriate. Odondi (1999) applied the Leslie age-structured population matrix on fish population and derived the parameters of the model for application in optimal fish harvesting.

Most of the cited works above are mainly deterministic in nature and therefore does not incorporate variability in key population parameters such as carrying capac-

ity or natural mortality. The approach used in these methods for parameter estimation has used classical likelihood-based statistical theory. The problem of computing likelihood function requires high dimensional integration because it is necessary to integrate over the unobserved process errors (Schmute, 1994). A tractable integral is obtained when the model is linear and has normally distributed errors (Meinhold and Singpurwalla 1983). But linearity and normality greatly restrict the realism and general applicability of the state-space model methodology. One way out of this problem would be to use penalized likelihood where process errors are treated as fixed parameters to be estimated. But this approach also has a problem in that using penalized likelihood to fit generalized linear models makes the estimates of fixed effects inconsistent when there are limited data per random effect (Lin and Breslow, 1996). The Bayesian state-space approach that uses the Markov Chain Monte Carlo (MCMC) techniques for sampling from joint posterior are very general in their application and do not rely on such assumptions as linearity or normality. Bayesian estimation allows for the inclusion of information from diverse sources through use of prior probabilities. The thrust of this study will therefore be to attempt to apply this approach in order to devise an appropriate statistical model for assessing stock in an aquatic population with application to a specific fish population.

### 1.3 Statement of the Problem

Fish is an important resource both as food (source of protein) and as a source of income to the family units and foreign exchange earner to the country. Fish stocks are renewable resources if correctly handled. If all mature fish are caught before spawning, then there will be no recruitment. Therefore, there should be sufficient <sup>numbers</sup> ~~amount~~ of mature fish in the stock at each time. The problem is to find the appropriate middle ground where it is possible to obtain good catches for a long time. It is necessary to estimate the size and productivity of fish stocks. In order to estimate the abundance and productivity of fish, it is necessary to have measurements which relate to the stock size in some way. It is rarely possible to count the abundance of a marine species, yet it is known that measurements are made of quantities which are only



indirectly related to the stock size. Therefore mathematical models are used to relate the measurements to the stock size and statistical techniques are then used to estimate the unknown parameters in such models.

It may also be necessary to make some statement about the potential yield. In such instances, attempts are made to investigate how sensitive results may be in terms of changes in assumptions of variability in the input data. It is also of importance to answer questions on how certain catches will affect the stock and yield potential in the future. A fish manager may be interested in knowing how a stock can be utilised in the long run or what will be the likely consequences of certain actions as regards catches and developments of the stock. An advisor may be asked to construct an entire management system, which could be of the form of a computer program which takes as input biological information and returns a quota. The design of the system must take into account all available information about the stock.

The desire to develop sound policies on fish management in terms of breeding and harvesting may be hampered by lack of knowledge on their dynamics. A good mechanism for monitoring fish stock and putting into consideration the fish characteristics such as sex, age, species, morphology and even predation effect can help one devise techniques that will optimize the fish breeding and harvesting. This study will therefore attempt to develop a statistical model for recruitment and development at various stages of life and use empirical data to estimate the parameters of the model.

#### **1.4 Aims and Objectives**

In this study, the aim is to develop an ecologically sustainable statistical model that can be applied to describe the dynamics of aquatic population through their life stages. The developed model will be applied to fish population. More specifically, the study will use empirical data to:

- (a) Determine the maintainability and attainability structure of a fish population
- (b) Apply the Markov Chain Monte Carlo (MCMC) method and the Bayesian

procedure to estimate the parameters of the derived model

- (c) Obtain the confidence bounds for the derived parameters, and
- (d) Simulate the behaviour of the model when external forecasted factors such as weather/climate and pollution are factored into the model.

### 1.5 Significance of the Study

The basic purpose of fish stock assessment is to provide advice on the optimum exploitation of aquatic living resources such as fish. Living resources are limited but renewable, and fish stock assessment may be described as the search for the exploitation level which in the long run gives the maximum yield in weight from the fishery. The model developed in the study will be an important management tool for sustainable fishery exploitation. The tool will be useful to researchers and scientists working on aquatic systems. It will also be an important analytical tool for fish policy planners. The new approach of using Bayesian techniques for estimating parameters of a population will be of great interest to statisticians as it will give new dimension in the application of Bayesian estimation as opposed to the traditional classical likelihood functional approach.

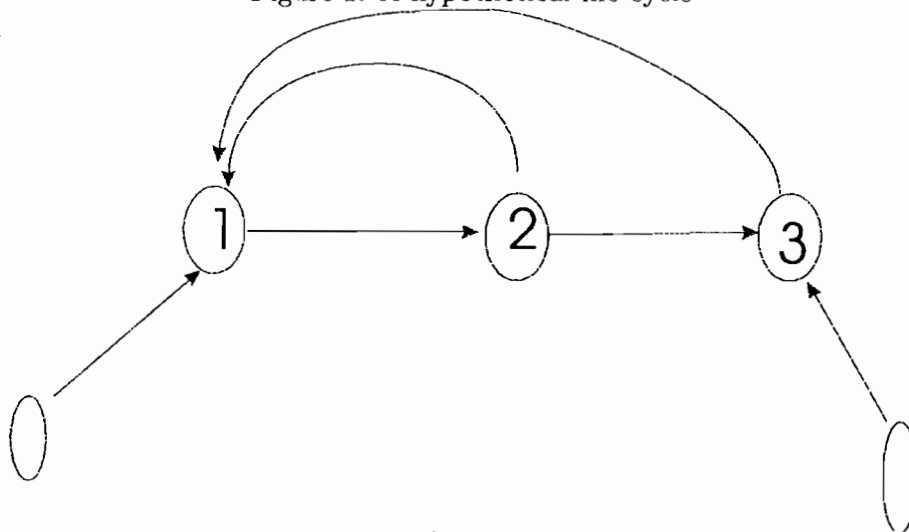
## 2 METHODOLOGY

### 2.1 The Basic Matrix Modeling Framework

Matrix models classify life cycles either by age or stage. In age-classified models, population dynamics are examined by dividing the continuous age variable into discrete age classes, each of same duration. Transitions among the stages are governed by vital rates that describe processes such as growth, fertility, survival or even the probability of becoming a breeder or obtaining a high quality territory. This model assumes the absence of density dependence and constancy of the age-specific vital rates, which in turn makes it less robust since age may not be the most appropriate basis for assessing population dynamics e.g. individuals of the same age may differ

in reproductive status. The stage-structured models can be used to model and analyze complex social dynamics or breeding systems as opposed to age structured. The study will attempt to apply both the age-structured and size-structured approaches for model evaluation. In the life cycle graph below, the nodes are translated as the

Figure 1: A hypothetical life cycle



stages or states of the system, the loops connecting the nodes represent the transition rates between the states. We let  $P_x$  to denote the transition rates indicating the survival rate from stage or age class  $x$  to  $x + 1$  and  $F_x$  to indicate the mean number of female offsprings produced by females of age  $x$ ,  $m_x$ . The adult survival is represented by an arc (or loop) that represents the probability of returning to the adult stage at the next census. The growth rate  $\lambda$  of the population is given by the dominant eigenvalue of the population matrix,  $A$ .

## 2.2 Data Collection and Analysis

The study will make use of primary and secondary data to be obtained from fisheries research institutions such as the Kenya Marine Fisheries Research Institute. If it will be deemed appropriate, the study will apply a suitable mark-recapture technique to sample some fish species of interests. The Markov Chain Monte Carlo (MCMC) technique and particularly the Bayesian approach shall be used extensively in deriving

the parameters of the modeled population. The associated software, the Bayesian Inference Using Gibbs Sampling (BUGS) software, shall be used for the analysis. Bayesian estimation makes use of prior information regarding the population under study. The software implements general Bayesian models using what is known as the Metropolis-Hastings within Gibbs sampling. This approach has been lauded as the only coherent statistical methodology for updating knowledge using the information contained in the data. This property enables the posterior from one analysis to be used as an induced prior in a subsequent analysis, thereby building and exploiting an accumulated base of knowledge.

### **2.3 Demographic Sensitivity, Elasticity and Fitness Analysis**

Demographic sensitivity means the sensitivity of the population growth rate  $\lambda$  to changes in particular demographic rates. It measures the impact on  $\lambda$  of changing the absolute magnitude of a particular vital rate (e.g. second year survival) relative to changing other vital rates. Different values of the population parameters such as fishing efforts and variable stock sizes shall be used to measure the sensitivity of the growth rate on those parameters. Elasticity analysis shall be made to assess the proportional sensitivity of  $\lambda$  to changes in a vital rate. Eigenvalue analysis of the projection matrix yields the sensitivities and elasticities. This will in turn lead to the fitness analysis where suitable rates of the population parameters are recommended based on specific desired objectives. The fitnesses are computed from the dominant eigenvalue of the reproductive outputs in the top row and survival values of 1.0 in the subdiagonal, with the dimensionality of the matrix equaling the life span of the individual.

## 3 MATRIX POPULATION MODELS

### 3.1 Leslie Matrix Approach

The method of Leslie is one of the most heavily used models in population ecology. This is a discrete time model of an age-structured population which describes development, mortality, and reproduction of organisms. The model has been used to solve (i) the rate of exponential growth (intrinsic rate of increase) of the population, and (ii) the proportion of each class in the stable age distribution.

**The Model Structure:** The model of Leslie (1945) describes three kinds of ecological processes, development (progress through the life cycle), age-specific mortality, and age-specific reproduction. The model is described by two equations;

$$N_{x+1,t+a} = N_{x,t}S_x \quad (1)$$

and

$$N_t = \sum_{x=0}^n N_{x,t}m_x = N_{0,t}m_0 + N_{1,t}m_1 + \dots + N_{n,t}m_n \quad (2)$$

where  $N_{x,t}$  is the number of organisms in age  $x$  at time  $t$  (age is measured in the same units as  $t$ ), and only females are usually considered,  $S_x$  is the survival of organisms in age interval from  $x$  to  $x + 1$ ; and  $m_x$  is the average number of female offsprings produced by one female in age interval from  $x$  to  $x + 1$  (mortality of parent and/or offspring organisms is included). Equation (1) represents the development and mortality of the organism, whereas equation (2) represents reproduction. The two equations are combined into one matrix equation

$$N_{t+1} = A.N_t \quad (3)$$

where  $N_t$  is the vector of age distribution in the population at time  $t$ , and  $A$  is the **transition matrix**. Equation (3) can be represented by

$$\begin{bmatrix} m_0 & m_1 & \dots & m_{n-1} & m_n \\ s_0 & 0 & \dots & 0 & 0 \\ 0 & s_1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & s_{n-1} & 0 \end{bmatrix} \begin{bmatrix} N_{0,t} \\ N_{1,t} \\ \vdots \\ N_{n-1,t} \\ N_{n,t} \end{bmatrix} = \begin{bmatrix} N_{0,t+1} = \sum N_{x,t+1}m_x \\ N_{1,t+1} = N_{0,t}S_0 \\ N_{2,t+1} = N_{1,t}S_1 \\ N_{3,t+1} = N_{2,t}S_2 \\ N_{4,t+1} = N_{3,t}S_3 \end{bmatrix} \quad (4)$$

or

$$N_{t+1} = AN_t$$

and generally

$$N_t = A^t N_0 \quad (5)$$

**Model Behaviour:** The model of Leslie exhibit two major characteristics namely, a few oscillations followed by an exponential growth that approaches a stable age distribution.

## 4 STOCK ESTIMATION

If recruitment in a stock is variable and if the species is a long-live one, it is necessary to investigate how year-classes in the stock develop. It is not enough to view the stock as an entity or to view the stock by length classes. It is natural that the fish population reduces as a function of age unless there is immigration into the area. One way, therefore, of stock estimation is based on the assumption that within a specific time period there is a constant fishing pressure and constant natural mortality. Denoting the number of fish a year-class at time  $t$  as  $N_t$ , the proportional change in numbers in a short time period,  $\Delta t$ , should be in proportion to the length of the time period,

$$\frac{\Delta N}{N} = -Z\Delta t, \quad (6)$$

where  $Z$  is some number. This leads immediately to the differential equation whose solution is of the form

$$N_t = e^{-Zt}N_0 \quad (7)$$

which gives the size of year-class at time  $t$ ,  $N_0$  is the size of the initial stock. The coefficient  $Z$  is the total mortality rate. This equation is modified as

$$N_{a+1,y+1} = e^{-Z_{ay}}N_{ay} \quad (8)$$

On a logarithmic scale there is a convenient relationship between the mortality rate and the changes in stock size expressed as

$$\ln N_{a+1,y+1} = \ln N_{ay} - Z_{ay} \quad (9)$$

The mortality rates here simply denotes the reduction in the stock on a logarithmic scale. It would be of interest to take a specific year-class and investigate how its biomass develops as the year-class grow older. Such an analysis gives an indication of rational utilization of the year class. If there are no catches from the year-class, then the mortality coefficient  $Z_{ay}$  simply denotes the natural mortality and is denoted  $M_{ay}$ . It is then possible to compute how many fish survive to the following year and thus an estimate is obtained of how many fish from this year-class will be alive at each moment in time. The counts can then be multiplied by the mean-weight at age to obtain the biomass at age.

## 5 MODELS FOR VARIABLE FISH RECRUITMENT

Not only must we allow the fish to grow and attain a good weight, we must also allow a sufficient number to spawn and thereby perpetuate the stock. If too many fish are caught before they have had an opportunity to contribute to later generations, then the stock is said to be experiencing recruitment overfishing. For most fish species, the adult females produce vast numbers of eggs during their life spans. Most of these eggs do not survive. For any population to remain stable over time, each adult female on average must produce exactly two offsprings (one male and one female) that survive and reproduce. If there were more than two successful offsprings per female, the population would increase in number. If there were less than two, the population would decrease. Most fish populations exhibit tremendous variability in the number of young fish that recruit annually. It is not unusual for the largest cohorts at the age of recruitment to be 100 times more abundant than smallest cohorts. Understanding recruitment variability is very important for the rational and successful management of any fishery. Recruitment must somehow depend on the size of the parent population i.e. if no fish, then no eggs, hence no recruits.

### 5.1 Ricker Spawner-Recruit model

In this model, a recruit is defined as a fish that survives to join the population of spawning adults. Spawning (maturity) can occur before or after the fish became available to fishery (Ricker, 1954). The Ricker spawner-recruit (or simply Ricker SR) model is defined by the differential equation

$$\frac{dE}{dt} = -M_E E \quad (10)$$

where

$$M_E = M + \beta S \quad (11)$$



and  $M$  is the density-independent rate of natural mortality,  $\beta$  is the density dependent term,  $S$  is the number of spawning adults, and  $E$  is the total number of eggs which then becomes larvae, juveniles and then adults. The term  $\beta S$ , the density dependent term, means that some of the instantaneous natural mortality is proportional to the abundance of the parent population. This form of mortality could arise from processes such as cannibalism, concentration of predators, or competition for spawning sites. The general solution of the differential equation is of the form

$$E(t) = kS e^{-(M+\beta S)t} \quad (12)$$

where  $kS = E_0$  at  $t = 0$ , and  $k$  is the average number of eggs laid per spawning adult. If all the fish in a cohort spawn at the same age and that they die after spawning, then  $E(t_s) = R$  is the recruitment where  $t_s$  is the age at spawning, and

$$R = aS e^{-bS} \quad (13)$$

where  $a = k e^{-mt_s}$  with units eggs/fish, and  $b = \beta t_s$  with unit 1/fish.

This is the Ricker spawner-recruit model and the term  $aS$  is the number of eggs that would survive to become spawners if the only source of natural mortality was  $M$  (a constant independent of the size of the parent population).

## 5.2 Beverton and Holt Spawner-Recruit Model

The model comes from the solution to the differential equation

$$\frac{dE}{dt} = -M_E - E \quad (14)$$

where  $M_E = M + \alpha E$ , and  $\alpha E$  is the density dependent rate of natural mortality,  $M$  is the density independent rate of natural mortality. The density dependent term is proportional to the cohort abundance, not to the parental abundance. This form of mortality could arise from processes such as competition for food or other scarce resources. The solution to equation (31) will be of the form

$$R = \frac{S}{C + dS} \quad (15)$$

where

$$C = \frac{1}{K} e^{Mt_s} \text{ and } d = \frac{\alpha}{M} (e^{Mt_s} - 1) \quad (16)$$

This is the Beverton and Holt SR model. If  $R = S$ , the spawning population will just replace itself with new recruits and the population will be at equilibrium from generation to the next. The equilibrium stock size is given by

$$S_{eq} = \frac{1 - C}{d} \quad (17)$$

## 6 HARVESTING AND STOCK RECRUIT MODELS

If fishing occurs over a short interval just prior to spawning, then the incoming recruits include progeny that spawn plus fish that would have spawned had they not been caught prior to spawning i.e. *Recruits = New spawners + catch*. If we adjust the size of our catch so that we harvest all recruits that are in excess of the replacement level,  $R = S$ , then the population will be maintained at a steady state with constant stock size from one generation to the next. The equilibrium catch per cohort is given by

$$C_{eq} = R(S) - S = S [a e^{-bS} - 1] \quad (18)$$

The maximum equilibrium catch occurs at that value of  $S$  for which

$$\frac{dC_{eq}}{dS} = \frac{d}{dS} (R(S) - S) = 0 \quad (19)$$

and satisfies

$$ae^{-bS}(1 - bS) = 1 \quad (20)$$

whose solution is of the form

$$\max C_{eq} = SC(ae^{-bSC} - 1) \quad (21)$$

## 7 SURPLUS PRODUCTION MODELS

Often in fisheries the main concern is with the surplus of births over deaths, the so called surplus production. The excess of births over deaths is considered to be surplus because it is not needed to maintain the population. There are different models that can be used to estimate the surplus production.

### 7.1 The Graham-Schaefer Surplus Production Model

The first application in fisheries of the logistic growth model was by Graham (1935) and Schaefer (1954). The model is now usually described as the Schaefer production model. The basic Graham-Schaefer model is defined by

$$\frac{dB}{dt} = rB \left[ 1 - \frac{B}{K} \right] - FB = \text{Natural Growth} - \text{Harvest} \quad (22)$$

If the population is at equilibrium, then

$$\frac{dB}{dt} = 0 = rB_e \left[ 1 - \frac{B_e}{K} \right] - F_e B_e \quad (23)$$

or

$$F_e B_e = Y_e = rB_e \left[ 1 - \frac{B_e}{K} \right] \quad (24)$$

where  $Y_e$  is the equilibrium yield and  $B_e$  is the equilibrium biomass.

## 7.2 Pella and Tomlinson's Generalized Surplus Production Model

One problem with the Graham-Schaefer model is that the maximum sustainable yield ( $MSY$ ) always occurs when the biomass is half the carrying capacity  $K$ . This is a direct consequence of the parabolic relationship between  $dB/dt$  and  $B$ , which in turn follows from the linear relationship between per capita productivity and population size. Pella and Tomlinson (1969), proposed an alteration to the model for latent productivity, which uncouples  $B_{msy}$  from  $K$ , given by

$$\begin{aligned} aB^n - bB \dots & \text{for } 0 < n < 1 \\ \frac{dB}{dt} = bB - aB^n \dots & \text{for } 1 < n. \end{aligned}$$

This is not a convenient formulation for this model. The values for  $MSY$ ,  $K$ , and  $B_{msy}$  all depend on parameter  $n$ , and you have to reverse the signs of parameters  $a$  and  $b$  depending on whether  $n$  is greater or less than one. Fletcher (1978) proposed alternative parameterization to avoid the problems mentioned above,

$$\frac{dB}{dt} = \gamma MSY \left[ \frac{B}{K} \right] - \gamma MSY \left[ \frac{B}{K} \right]^n$$

with

$$\gamma = \frac{\frac{n}{n^{n-1}}}{n-1}. \quad (25)$$

The parameter  $\gamma$  (gamma), which is a pure number, automatically changes sign as  $n$  increases through one. Parameter  $n$  controls the location of  $B_{msy}$ . Differentiating the above equation again and equating to zero gives the value of  $B_{msy}$  as

$$B_{MSY} = K n \frac{-1}{n-1} \quad (26)$$

When  $n = 2$ , the model is equivalent to the Graham-Schaefer model. When  $n > 2$ ,  $B_{MSY} > \frac{K}{2}$  and maximum productivity is closer to  $K$ . When  $n < 2$ , and maximum productivity is closer to zero. When  $n = 1$ , the system reduces to the so called "exponential" surplus-production model of Fox (1970),

$$\frac{dB}{dt} = -e.MSY.\frac{B}{K} \ln\left(\frac{B}{K}\right) \quad (27)$$

These surplus production models do not specify the biological processes responsible for the curvature in the graph of  $dB/dt$  versus  $B$ .

## 8 MARKOV CHAIN MONTE CARLO METHODS

### 8.1 Ordinary Monte Carlo Computations

The Markov Chain Monte Carlo (MCMC) methods have had profound influence on statistics over the past dozen years especially in Bayesian Inference. The implementation of Ordinary Monte Carlo sampling is rarely feasible in practice, except for the types of rather simplistic or contrived problems. Nevertheless, the underlying ideas transfer quite smoothly to MCMC, with random samples replaced by independent samples from a Markov chain.

Suppose we let  $X$  denote a random quantity which could be a random vector or a multi-way contingency table. Then  $X$  have many components and some of the

components may be discrete and others may be continuous. Further take  $X$  as single random variable with finite but complicated sample space. Let  $\{\pi(x); x \in S\}$  denote the probability distribution of  $X$ , where  $S$  is the corresponding minimal sample space, i.e.  $S = \{x : \pi(x) > 0\}$ . We assume that  $\pi(\cdot)$  is known up to scale, so that

$$\pi(x) = h(x)/c, x \in S. \quad (28)$$

where  $h(\cdot)$  is completely specified and

$$c = \sum_{x \in S} h(x) \quad (29)$$

is usually not known and the space  $S$  is too large for  $c$  to be calculated directly. The purpose is to compute the expectation of some function  $g$  under  $\pi$  i.e.

$$E_{\pi}g = \sum_{x \in S} g(x)\pi(x) \quad (30)$$

for any relevant  $g$ . This summation cannot be carried out directly (even in the rare event that  $c$  is known). This equation includes the probability of any particular event concerning  $X$ . For any relevant subset  $B$  of the minimal sample space  $S$ ,

$$\Pr(X \in B) = \sum_{x \in S} 1[x \in B]\pi(x), \quad (31)$$

where  $1[\cdot]$  is the indicator function i.e.  $1[x \in B] = 1$  if the outcome  $x$  implies that the event occurs and  $1[x \in B] = 0$  otherwise.

Let  $x^{(1)}, x^{(2)}, \dots$  be random variables from the target distribution  $\pi$  corresponding to independent and i.i.d random variables  $X^{(1)}, X^{(2)}, \dots$ . If we produce  $m$  such draws  $x^{(1)}, \dots, x^{(m)}$ , then the estimate of  $E_{\pi}g$  is

$$\bar{g} = \frac{1}{m} \sum_{t=1}^m g(x^{(t)}). \quad (32)$$

## 8.2 Bayesian Computation

Let  $x$  denote an unknown scalar (parameter) in a finite parameter space  $S$  and suppose that  $\{\rho(x) : x \in S\}$  is a prior probability distribution representing the initial belief about the true value of  $x$ . Let  $y$  denote relevant data, with corresponding known likelihood  $L(y|x)$ , so that the posterior probability distribution for  $x$  given  $y$  is

$$\pi(x|y) \propto L(y|x)\rho(x), \quad x \in S. \quad (33)$$

By equations (28) and (29), we obtain

$$h(x) \propto L(y|x)\rho(x) \quad (34)$$

In Bayesian paradigm, inferences are conditional on the fixed data  $y$ . If  $x^{(1)}, \dots, x^{(m)}$  is a large random sample from  $\pi(x|y)$  and for fixed  $y$ , then with the appropriate choices of  $g$ , equation (32) can be used to closely approximate the posterior mean and variance, and more importantly to evaluate the posterior probabilities concerning the parameter  $x$  and to construct corresponding credible intervals.

## 8.3 Markov Chains and Monte Carlo calculations

Let  $X^{(0)}, X^{(1)}, \dots$  be a Markov Chain with transition probability matrix  $P$  and state space  $S$  and define  $p^{(0)}$  to be the row vector representing the distribution of the initial state  $x^{(0)}$ . The marginal distribution of  $X^{(t)}$  is given by

$$p^{(t)} = p^{(0)}P^t, \quad t = 0, 1, \dots, \quad (35)$$

and that if  $\pi$  is a probability distribution vector satisfying general balance  $\pi P = \pi$ , then  $\pi$  is called the stationary distribution for  $P$ , i.e.  $p^{(t)} \rightarrow \pi$  as  $t \rightarrow \infty$  irrespective of  $p^{(0)}$ . This implies that by the ergodic theorem for Markov chains equation (32) converges to  $E_\pi g$  as  $m \rightarrow \infty$ .

In Bayesian approach, analysis corresponds to determining the posterior distribution of unknowns of interest. The BUGS software is a Markov chain Monte-Carlo tool that samples numerically from the joint posterior of the unknown quantities. In Bayesian decision analysis, the posterior expected values and distributions of indicators are calculated through the MCMC simulation.



## 9 LIST OF REFERENCES

1. Albrecht, F., H. Gatzke, A. Haddad, and N. Wax (1974). The dynamics of two interacting populations. *Journal of Mathematical Analysis and Applications*, 46:658-670
2. Bartlett, M.S.(1960). *Stochastic Models in Ecology and Epidemiology*. Methuen, London.
3. Beddington, J.R.(1973). Optimum age-specific harvesting of a population. *Biometrics*, 24:1-10.
4. Beddington, J.R.(1984). Management under uncertainty, in *Exploitation of Marine Communities*, R.M. May (ed.) 222-244, Springer-Verlag, Berlin.
5. Bernadelli, H.(1941). Population Waves. *Journal of the Burma Research Society*. 31:1-18
6. Burgman, M.A., S. Ferson, and H.R. Ackakaya (1993). *Risk Assessment in Conservation Biology*. Chapman Hall. London.
7. Caswell, H. (1989). *Matrix Population Models*. Sinauer Associates, Inc., Massachusetts.
8. Cole, L.C. (1958). *Sketches of General and Comparative Demography*. Cold Spring Harbor Symposia on Quantitative Biology. 22:1-15.
9. Doubleday, W.G. (1975). Harvesting in Matrix Population Models. *Biometrics*, 31:189-200.
10. Fletcher, (1978). On the restructuring of the Pella-Tomlinson system. *Fish. Bull.* 76:515-521.
11. Fox, W.W. Jr.,(1970). An exponential surplus-yield model for optimizing exploited fish populations. *Trans.Am.Fish.Soc.*, 99:80-88
12. Gause, G.F., (1934). *The struggle for existence*. Dover Publications, Inc. New York.

13. Gillian, J.F. , and D.F. Fraser (1987). Habitat selection under predation hazard: test of a model with foraging minnows. *Ecology*. 68:1856-62.
14. Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality. *Philosophical Transactions*. 115:513-585.
15. Graham, M., (1935). Modern theory of exploiting a fishery and application to North Sea trawling. *J.Cons.CIEM*, 10(3):264-274
16. Hastings, A., C.L. Hom, S. Ellner, P. Turchin, and H.C.J. Godfray (1993). Chaos in ecology: is Mother Nature a strange attractor?. *Annual Review of Ecology and Systematics*. 34:1-33.
17. Hutchinson, G.E. (1978). *An Introduction to Population Ecology*. Yale University Press, New Haven.
18. Kerr, S.R. (1982). Niche theory and fisheries ecology. *Trans. Am. Fish Soc.* 109:254-257.
19. Kingsland, S. (1985). *Modeling Nature*. University of Chicago Press, Chicago.
20. Kolmogorov, A.. (1936). Sulla teoria di Volterra della lotta per l'esistenza, *Gi. Inst. Attuari*. 7:74-80.
21. Koztitzin, V.T. (1939). *Mathematical Biology*. George G. Harrap and company Ltd. London.
22. Lefkovitch, L.P. (1965). The study of population growth in organisms grouped by stages. *Biometrics*. 21:1-15.
23. Leslie, P.H. (1945). On the use of matrices in certain population mathematics. *Biometrika*. 33:183-212.
24. ....(1948). Some further notes on the use of matrices in population mathematics. *Biometrika*. 35:213-235.
25. Lewis, E.G. (1942). On the generation and growth of a population. *Sankya: The Indian Journal of Statistics*. 6:93-96.

26. Lin, X., and N.E. Breslow (1996). Bias correction in generalized linear mixed models with multiple components of dispersion. *J. Amer. Stat. Assoc.* 91: 1007-1016.
27. Lotka, A.J. (1924). *Elements of Physical Biology*. Williams and Wilkins, Baltimore.
28. .... (1939). A contribution to the theory of self-renewing aggregates, with special reference to industrial replacement. *Annals of Mathematical Statistics*. 10:1-25.
29. Malthus, T.R. (1798). *An essay on the principle of population and a summary view of the principle of population*. Penguin, Harmondsworth, England.
30. May, R.M. (1972). Limit cycles in predator-prey communities. *Science*. 177:900-902.
31. .... (1973). *Stability and complexity in model ecosystems*. Princeton, New Jersey, USA.
32. May, R.M., and G.F. Oster (1976). Bifurcations and dynamic complexity in simply ecological models. *American Naturalist*. 110:573-599.
33. Meinhold, R.J., and N.D. Singpurwalla (1983). Understanding the Kalman filter. *Amer. Stat.* 37: 123-127.
34. Odoni, L.T. (1999). *Optimal Fish Farming*. M.Phil. Thesis submitted to Mathematics Department, Moi University, Kenya.
35. Oster, G.F. (1975). Review of "Ecological Stability" (M.B. Usher and M.H. Williamson, Eds.), *Ecology*. 56:1462.
36. Pauly, D. (1980). On the relationship between natural mortality, growth parameters and mean environmental temperature in 175 fish stock. *Rapp. P.V. Reun. Cons. Int. Explor. Mer.* 39:175-192.
37. Rescigno, A., and I.W. Richardson (1967). On the competitive excluding principle. *Bulletin of Mathematical Biophysics*. 27:85-90.

38. Ricker (1954). Stock and recruitment. In: Key Papers on Fish Populations, p.151-225.
39. Sauer, J.P., and N.A. Slade (1987). Underground demography: Is body mass a better categorical variable than age? *Ecology*. 68:624-650.
40. Schaefer, M., (1954). Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Bull.I-ATTC/Bol. CIAT*, 1(2):27-56
41. Schnute, J.T., (1994). A general framework for developing sequential fisheries models. *Can J. Fish. Aquat. Sci.* 51: 1676-1688.
42. Smith, P.E. (1981). Fisheries on coastal pelagic schooling fish. In *Marine Fish Larvae: Morphology Ecology, and Relation to Fisheries*. R. Lasker (ed.) pp 1-32. University of Washington Press, Seattle.
43. Strong, D.R. (1984). Density-vague ecology and liberal population regulation in insects. In *A New Ecology: Novel Approaches to Interactive Systems*, P.W. Price, G.N. Slobodchikoff and W.S. Gaud (eds.). pp 313-329. Wiley, New York.
44. Turchin, P. (1995). Population regulation: old arguments and a new synthesis. Pages 19-40 in N. Cappuccino and P.W. Price editors - *Population dynamics: new approaches and synthesis*. Academic Press, San Diego, California, USA.
45. Usher, M.B. (1966). A matrix approach to the management of renewable resources, with reference to selection forests. *J. Appl. Ecol.* 3:355-387.
46. .... (1969). A matrix model for forest management. *Biometrics*. 25:309-315.
47. Verhulst, F.P. (1838). Notice sur la loi que la population suite dand son Acroissement. *Correspondance Mathathematique et Physique*. 10:113-121.
48. Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature*. 118:558-560.
49. Werner, E.E. (1982). Niche theory in fisheries ecology. *trans. Am. Fish Soc.* 109:257-260.

## 10. TIME PLAN

Activity	Duration
<b>YEAR 1</b>	
Literature and software Acquisition	3 Months
Literature Review	3 Months
Theoretical Model building	3 Months
Review of Matrix Population models	3 Months
<b>YEAR 2</b>	
Review existing models for stock estimation	3 Months
Data collection, cleaning and validation	3 Months
Develop Bayesian-based model for stock assessment	3 Months
Model simulation	3 Months
<b>YEAR 3</b>	
Model Evaluation and validation	3 Months
Workshop and seminar presentations	3 Months
Draft Thesis compilation	3 months
Corrections and Final submission of thesis	3 Months
<b>Total</b>	<b>36 Months</b>

## 11. BUDGET

Item	Approximate Amount (Kshs)
<b>Year One</b>	
<b>Computer and Accessories</b>	
a. Laptop Computer	200,000
b. Printer Cartridge ? 500 for Each ?	32,000
c. Diskettes (5 boxes)	2,500
d. UPS and Printer (Laser Jet)	50,000
e. Software Acquisition (BUGS software)	150,000
<b>Stationary</b>	
a. Printing paper (5 reams)	2,500
b. Fullscaps (5 reams)	1,500
c. Pens (4 dozens)	400
<b>Sub-Total</b>	<b>341,300</b>
<b>Year Two</b>	
<b>Photocopying and Acquisition of Journals</b>	
a. Locally available journals (2,000 pages)	6,000
b. Journals not locally available (15 copies)	40,000
c. Other materials	10,000
<b>Stationary</b>	
a. Printing paper (5 reams)	2,500
b. Fullscaps (5 reams)	1,500
<b>Data Collection</b>	100,000
<b>Conference and seminars</b>	130,000
<b>Sub-Total</b>	<b>234,000</b>
<b>Year Three</b>	
Conferences and Seminars	120,000
Thesis Preparation and Binding	50,000
<b>Sub-Total</b>	<b>300,000</b>
<b>GRAND TOTAL</b>	<b>745,300</b>

