STOCHASTIC MODELING OF BAMBOO POPULATION GROWTH AND OPTIMAL HARVESTING

BY

ARORI WILFRED OMWANSA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN APPLIED STATISTICS

SCHOOL OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

MASENO UNIVERSITY

DECLARATION

I declare that the work in this thesis has not been previously presented for a degree award at Maseno University, or any other University. The work reported herein has been carried out by me and all external sources of information have been acknowledged by means of references.

Signature: _____   Date: _____

Arori Wilfred Omwansa,
PG/PHD/051/2008

This thesis has been submitted with our approval as the university supervisors

Signature: _____   Date: _____

Prof. Fredrick Onyango,
School of Mathematics, Statistics and Actuarial Science,
Maseno University, Kenya.

Signature: _____   Date: _____

Prof. Jacob K. Kibwage,
The Cooperative University of Kenya.

ACKNOWLEDGEMENT

DEDICATION

To my son Saint Xavier who was an inspiration as I made effort to complete this work.

ABSTRACT

Population growth and harvest modeling is an active area of current research. There has been an effort to move from deterministic Ordinary Differential Equations (ODE) to Stochastic Differential Equations (SDE) modeling. Moreover, the latter is most realistic in describing life systems that are often perturbed by unpredictable environmental activity. Bamboo growth and harvest modeling was motivated by the "Tobacco to Bamboo" (TTB) Project where farmers in selected sections of Homabay and Migori Counties in Kenya were persuaded to plant bamboo instead of tobacco. This was met with pessimism due to the lengthy wait, at least three years, before harvesting. They also needed to know the expected income compared to the tobacco income they used to earn. This study therefore sought to explore suitable models that could be used to determine optimal expected sustainable bamboo yield. In view of this, data from the TTB project was analyzed to determine parameters including population growth rate r, carrying capacity K and population size at time t, N_t. ODEs and SDEs were used in modeling equilibrium populations and maximum sustainable yield. SDEs were solved using Itô calculus and associated Fokker–Planck equations. The Monte-Carlo simulation procedure was used to construct population trajectories under various model parameter values. A stochastic model with both growth rate and harvest parameters coupled with white noise and a three year delayed continuous harvest proportional to population size was developed. This was found to be most suitable since it ensures maximum mean sustainable yield without the risk of extinction as long as noise was kept at low levels. The model may not only be applied in bamboo harvesting strategies but also other renewable resources that have similar population dynamics.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

ANOVA   Analysis of Variance

APSIM   Agricultural Production Systems Simulator

APSRU   Agricultural Production Systems Research Unit

CGHB   County Government of Homa Bay

CGM   County Government of Migori

Dbh   Diameter at breast height.

GIS   Geographic Information Systems

GOK   Government of Kenya

HVol   Harvest Volume (m^3/acre)

IDRC   International Development Research Center

INBAR   International Network for Bamboo and Rattan.

KEFRI   Kenya Forestry Research Institute

RGR   Relative growth rate

UNIDO   United Nations Industrial Development Organization

LHS   Left hand side.

RHS   Right hand side.

GOF   Goodness of fit

SLM   Stochastic logistic model

SPSS   Statistical package for social sciences

RNG   Random number generator

Mathematical notations

Anoc   Average number of culms (stems) per clump (bush).

K Carrying capacity.

$E(.)$ Expected value

$P(.)$ Probability density.

λ Proportion of mature culms demanded.

q Probability of finding a mature culm in a clump.

N_t Population size at time t.

r Intrinsic growth rate.

N_{eq} Equilibrium population.

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Population growth and decay are interesting dynamics for population modeling experts. There are a series of factors that affect population dynamics. These include temperature, rainfall, introduction of a new predator, congestion of the ecosystem due to uncontrolled population growth, human activity that could be pro or against the population in question, natural calamities including earthquakes, fires, landslides, floods and drought among others.

Some of these factors are controllable so that the system is easily represented by some precise and determinable variables. The reality is different though. In most cases it turns out that the dynamic variables are often perturbed by uncontrollable deviations that can only be explained by chance.

1.1   About bamboo

Bamboo is a member of the grass family, distinguished from other grasses by its woody stems, branched growth and often by large size. They are easy to grow, but some may take a lot of management to control unwanted spread. Worldwide, approximately 87 genera and over 1,500 species of bamboo exist [63, 7] and among these nearly 62 are found in the Philippines [79]. They range in size from dwarfs that are only 12 inches tall to giant types that tower up to 60 feet tall. A few bamboos have variegated foliage and some have very attractive canes of rose, burgundy, yellow, orange, gray or black. Although many types of bamboo are tropical there are also many that grow in temperate regions. The selection of desirable types of bamboo is much greater for southern gardeners than in the north where winter is the limiting factor. They are found growing in many parts of the world but are most common in south-east Asia, China and Japan.

Bamboo is widely distributed, mainly in the tropical, subtropical and mild temperate zones of the world, with the tropical belt having the maximum number of bamboo species.

Interest in the use of bamboo in the field of construction is growing around the world. This can be attributable to its good engineering properties as well as the fact that it is a renewable resource. There are several different bamboo types, with a great deal of variation in microstructure and mechanical properties. In the Brazilian North-eastern, *bambusa vulgaris* is the most common type. However, its durability and mechanical properties are usually inferior to many other species such as *dendrocalamus giganteus* and *Guadua angustifolia*.

Intensified interest on bamboos has resulted in their emergence as potentially the most important non-timber forest resource to replace wood. Their strength, straightness, lightness combined with extraordinary hardness, range in size, abundance, easy propagation, and short vegetation period to attain maturity, make them suitable for a variety of purposes and uses [36]. Furthermore, bamboo as a woody plant is uniquely suited to agro forestry. Some of the many uses of bamboo in agro forestry are: Intercropping; Riparian vegetation filter; Constructed wetlands; living screens; and perm culture [34].

The flowering habits of bamboo are not well understood. Some bloom every year while others bloom only once in fifty to one hundred years and then die depending on the variety and climate. Some species have never been known to bloom. When a bamboo shoot flowers, the event usually signals the end of its life cycle. *Melocanna baccifera* blossoms en masse. All bamboo stocks of the same flower-type blossom at the same time every 50 to 60 years, regardless of geographical location, and consequently die, leading in many cases to the seasonal starvation of animal creatures that normally feed on the bamboo plant. Though there are many theories as to what the reason behind bamboo mass blossoming is, the exact answer remains unknown.

Bamboo is the fastest growing woody plant in the world [51]. In a single day, a bamboo plant can grow an average of up to 60 centimeters (about 23.6 inches), and certain species have been measured to grow as much as 121 centimeters (about 47.6 inches) in one 24-hour period. It grows most abundantly in the East and Southeast Asian regions of the world, but it can also be found in Northern Australia, India, sub-Saharan Africa and the tropical regions of the Americas.

Bamboo has become a popular eco-friendly alternative to using hardwood as building material. Because of its rapid growth and short life-cycle, bamboo is one of the most renewable natural resources in the world, and it can be harvested on a regular basis without causing significant damage to its surrounding ecosystem.

1.2   Bamboo species

The accurate number of species is not accurately known. As earlier reported, 87 genera and over 1,500 species of bamboo exist [63, 7]. The American Bamboo Society [7] enlisted various species in their database. We are interested in two common species, *bambusa vulgaris* and *Dendrocalamus giganteus*. The two species were introduced in experimental forms described in Chapter Three.

1.2.1   *Bambusa vulgaris*

b. vulgaris is a medium sized clumping bamboo with pachymorph rhizomes. This species is, as its Latin name suggests, an ordinary bamboo in the sense that it is very common and widespread. It is cultivated throughout the tropics and is easily propagated by traditional means including culm cuttings.

There are several varieties within the species [72], the most common being *b. vulgaris 'vulgaris'*, which has culms that are bright glossy green in color . The variety *b. vulgaris* 'Vitatta' (also known as *b. vulgaris* var. *striata*) is distinguished by its glabrous, shiny, yellow culms with light green stripes. A variety with green culms and yellow stripes also exists, but is less common. The striped varieties are mainly used as ornamentals.

b. vulgaris is a multipurpose bamboo with culms of 8 to 20m in length. The diameter of culms is 5-10 cm with a wall thickness ranging between 7-15 mm [72]. It is used for paper-making, scaffolding, fencing, construction, poles, and handicrafts. The shoots of the species are edible, and consumed throughout south-east Asia. [48] studied Growth and development of *b. vulgaris schrad* planted on the coastal homesteads of Bangladesh. They reported culm height of 16.3 m and diameter of 5.29 cm with an average of 15.44 culms per clump at age 3 years and 21.2 culms at age 4 years.

b. vulgaris grows wild in the warmer parts of the country up to an altitude of 1,500 m, thriving best near river banks [50]. This widely cultivated species grows 6-18 m in height and 5-10 cm in diameter. The most commonly grown variety, the *striata* bears characteristic green streaks on a yellow background. This bamboo grows very fast and can attain a height of 9 m within 3 months. It is a highly versatile bamboo and is extensively used. The culms are relatively soft and poises long fibers, making them a valuable source of paper pulp. Ambika and Rajapogal [5] recently studied the potential of *b. vulgaris* as medicine for human use. *b. vulgaris* has perpetually remained vegetative with no flowering reported.

1.2.2 *Dendrocalamus giganteus*

d. giganteus is a tight clumping bamboo with pachymorph rhizomes. It is commonly known as Giant Bamboo. Originally from South East Asia, it has been introduced in numerous countries including India, Bangladesh and (southern) China. It is also cultivated in subtropical regions of the United States and in Australia. It grows naturally in humid tropical highlands, up to 1200 m above sea level and grows successfully in tropical lowlands on rich alluvial soils. Known to be one of the largest bamboo species in existence, its culms extend to heights of 35 m and with a diameter of up to 25 cm [72] and walls of 25 mm thick. Young shoots of the species are distinguished by their purple color. The full extension of the culm is achieved within three and a half months. The full giant size of the culm is attained only after 10 to 15 years. Thinning of the clump may begin 4 years after planting, removing the oldest culms and leaving space for the development for new shoots. 7 to 8 years after planting large 3 year old culms may be harvested. New shoots annually increase in size until the plant reaches maturity. The large culms of the species have many uses including construction, scaffolding, water pipes, buckets and paper pulp. The young shoots are edible and are distinguished by their creamy texture.

dendrocalamus giganteus (Giant Bamboo), native to Burma, may be the largest in the world. Its leaves can grow almost two feet long and four inches wide [50]. Bamboo is a common term for a large number of giant grasses that include many different species and varieties. There are two main types of bamboo. Runner types send out underground stems to varying distances and send up vertical shoots. These will grow in large thickets or groves if left alone. Runners are mainly found in temperate regions. Clump bamboos have underground stems that sprout vertical shoots much closer to their parent plants growing slowly outward. *D. longispathus* grows widely in parts of west Bengal, Assam and other areas of Eastern India, usually along streams. It can grow in an elevation of 4,000 m. the culms generally attain a height of 10-18 ft and a diameter of 6-10cm. It is one of the 15 economically important species of India that are recommended in the manufacture of paper.

1.3 Propagation and Growth of Bamboo

Bamboo propagation is one of the key links of bamboo production. It is used to obtain a great deal of fine seedlings for planting. Bamboo propagation can be classified into two kinds, one is sexual and another is asexual. Under each kind there are many sub-kinds for propagation. There are different propagation methods for different bamboo species.

Thus one should select a proper way of propagating them according to species specific requirements. Micro propagation of bamboos will be useful even with seed/seedling tissue as it would augment the supply of planting material. However, it would be ideal to perform micro-propagation with adult tissues as it would help in multiplying superior bamboo clumps on a large scale.

1.4   Harvesting of Bamboo

Depending upon the intensity of Bamboo occurrence, the following three categories are identified:

1. Scattered bamboo: With 50 clumps or less per hectare.

2. Dense bamboo: More than 50 clumps but less than 100 clumps per hectare.

3. Pure bamboo: More than 100 clumps per hectare.

The site quality indicates the capacity of site to produce bamboo culms of a given size. The best quality is indicated as an average height of the culm over 7 meters and average diameter at breast height of over 4 cm and in case of *bambusa arundinacea* an average diameter of 5 cm at breast height.

The production of new culms in the clump depends on the vigour of the underground rhizome, its growth activity and stored food in the previous growing season. Therefore, each clump must have sufficient number of green culms with adequate foliage to manufacture food and store in the underground rhizome. Repeated working will reduce the productive capacity of the clumps. Removal of congestion in the clumps, ensuring protection of underground rhizome from rotting, fire and other damages increases productive capacity of the clumps.

The best silvicultural system which satisfies all the requirements for maximum productive capacity of the clumps is selective felling of mature culms combined with tending of clumps to reduce congestion.

A three year felling cycle is found quite satisfactory to keep up the productivity. Therefore, the same has been prescribed in most of the working plans.

To provide congenial conditions for healthy growth of clump and culms, a 3 month rest period is prescribed from 1st July to 30th September every year [64]. This helps to protect

the rhizome from disturbance and prevents damage to tender bamboo shoots during the monsoon period.

The following felling rules apply [64]:

1. All current year culms should be retained.

2. A minimum of 6 culms over one year old, spaced uniformly all over the clump should be retained. In case of big clumps, proportionately more mature culms should be retained.

3. Culms should be felled between the first and second internodes from the ground level.

4. Felling cut should be given using a sharp instrument to avoid splitting and damage of culms.

5. All dead and dry bamboo and high cuts should be removed first.

6. Lopping of bamboo for feeding the livestock or otherwise is prohibited.

7. Clumps having 6 or less than 6 mature culms (more than one year old) in each clump will not be cut. Only broken, dead, dry or badly damaged and twisted bamboos are removed.

8. Rhizomes should not be damaged nor the clump uprooted.

9. All cut debris shall be removed at least one meter away from the periphery of each worked clump.

10. As far as possible in the year following working, the coupe should be closed from grazing.

11. Bamboo forests should be rigidly protected from fires especially during the year of working and the year following it.

12. When flowering is sporadic, all flowered clumps within the coupe under working and which have shed their seed, should be clear felled.

13. In case of gregarious flowering all clumps which have flowered, irrespective of felling cycle are clear felled, after seed is shed. The felled material should be removed expeditiously. These areas are strictly prohibited for grazing and are rigidly fire protected at least for 5 years to help establishment and growth of young regeneration. In addition, digging of cattle proof trenches wherever needed, should be taken up.

14. The congested bamboo clumps should be clear felled.

Unlike hardwoods, bamboo plants achieve their full height and girth in just one growing season, which lasts between three and four months. Bamboo shoots spend the first year stretching upwards, after which they begin to dry and harden and begin sprouting branches and leaves. After their second year, the shoots continue to harden even more, shedding their young sheath layers and becoming fully mature bamboo plants. After 5 to 8 years, bamboo shoots begin to decay and die, partly as a result of fungus and mold. Because of their short life-cycle, bamboo is best harvested between its 3rd and 7th years.

1.5 Mathematical modeling

Modeling population dynamics continues to be an area of interest to mathematicians and biologists. Growth and harvest estimation are key mathematical problems that arise in many real life situations. Methods of accomplishing this major goal have evolved over time, from deterministic equations that use differential calculus to more complex systems that use stochastic differential equations. The main reason for this shift is understandably the need to be realistic in estimates. Deterministic models assume that processes produce accurate results. However, in reality, the solutions are approximate. This gives the stochastic approach more credibility since solutions are assumed to vary each time we compute them.

Mathematical models have been applied in many areas of population studies. For example, there is a good amount of literature in modeling fish populations using the logistic equation [52, 83]. Moreover, various fishing strategies for mean sustainable yield (MSY) are documented.

1.6 The logistic model

The logistic model has been long used in ecological modeling due to its simplicity and effectiveness in portraying the growth of a population. The logistic equation has been widely studied and applied in population and ecological modeling. It describes population growth using a self-limitation term R which serves as a correction to the unlimited growth of the Malthusian model; the term is commonly referred to as the carrying capacity. The classic logistic (or Verhulst's) equation was suggested by Pierre Vehuslt [95] and is given

by:

$$\frac{dN_t}{dt} = rN_t\left[1 - \frac{N_t}{K}\right] \tag{1.1}$$

where N_t is the number of individuals at time t, r the intrinsic growth rate and K is the maximum number of individuals that the environment can support. Equation (1.1) is autonomous as the parameters r and K are constant. Its nature is such that the properties of the solutions, for all strictly positive initial conditions, approach the constant value of the carrying capacity, K, as time, t, tends to infinity (see Figure 1.1)). It is not often realistic to assume that the carrying capacity is constant, especially in modeling populations. Thus, many studies have discussed an alternative approach by investigating the importance of time-dependent carrying capacities, $K = K_t$ embedded in a non-autonomous logistic equation [31, 78, 41, 42, 43].
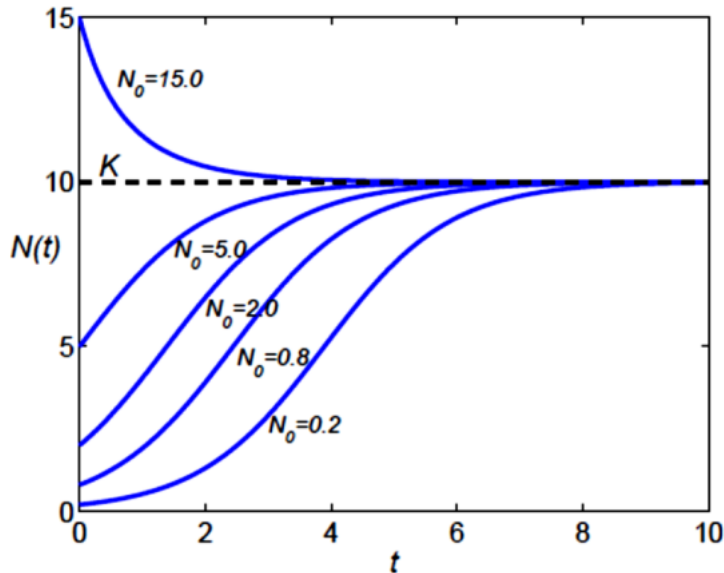
Figure 1.1: Logistic growth curves with $K = 10$ for various initial conditions N_0

1.7 Stochastic logistic model

A close observation on any population processes reveals that they are non-linear and stochastic in nature. In recent times a great deal of research work has been pursued to elucidate the role of nonlinearity and stochasticity in the evolution of the processes. Most of the work incorporating nonlinearities and stochasticity is either theoretical or qualitative in description. We have been motivated to make an empirical attempt to study stochastic logistic models of bamboo population growth. The stochastic logistic model can take many forms depending on which term in the system is considered volatile.

A general case where the growth rate is stochastic is given by:

$$dN_t = rN_t \left(1 - \frac{N_t}{K} \right) dt + \sigma N_t dB_t \qquad (1.2)$$

where σ is the diffusion coefficient which measures the size of fluctuations. B_t is the Brownian process with mean Zero and constant variance t. Equation (1.2) cannot be solved using the usual Riemann integration. This calls for Itô calculus and the Fokker-Planck equation. In our work, we move from the ordinary differential equations in the deterministic domain to stochastic differential equations under a noisy environment.

1.8   Statement of the Problem

Population growth and harvesting dynamics has attracted immense attention of researchers and mathematical population modelers in recent times. Furthermore, existing models such as Agricultural Production Systems Simulator (APSIM), [11] are crop specific. They are time series based and assume constant environmental and human effects. Further, none of them has incorporated bamboo as one of the possible input crops.

Various population models have been presented with variations dependent on the characteristics of the population under study. Most of these models are mainly based on the classical Malthusian theory of population growth, together with the Lokta-Volttera population interaction models. These models have their shortcomings considering the variability of growth dynamics from one population to another. They do not take into account the unpredictability of the environment in which the population under study is resident.

Our interest is in modeling Bamboo population growth and harvesting dynamics. We not only use the already established modeling techniques but also suggest introduction of effects of environment and human interferences commonly called white noise. This is done by invoking Brownian motion. This leads us to stochastic differential equations tailored to suit Bamboo populations. Other than deriving the stochastic differential equations under various harvesting strategies, we involve Monte-Carlo simulation techniques to determine the paths of Bamboo populations for each strategy. Further, a comparison of the paths under steady state and ruin scenarios are studied. Finally, the best strategy is supported for application by Bamboo farmers.

1.9   Objectives of the Study

1.9.1   General objective

The main objective is to model bamboo population growth and optimal harvesting strategies under a stochastic environment usig determinitic and stochastic differential equations.

1.9.2   Specific objectives

The specific objectives include;

1. Determining bamboo population growth parameters from experimental data.

2. Determining the most suitable growth model among a range of possible models.

3. Constructing deterministic bamboo population growth and harvesting models suitable to bamboo characteristics and determining optimum harvest.

4. Constructing a stochastic bamboo population growth and harvesting model and determining the optimum harvest.

5. Determining the possibility of ruin and time to extinction of bamboo population under deterministic and stochastic environments.

1.10   Significance of the study

Modeling parameters as a means of forecasting harvest is no doubt, a milestone in agricultural economics and planning. Suppose that a farmer, before planting a particular crop, say bamboo, already knows approximately how much he could earn from a given area of land. This is not only morale boosting but also valuable in early planning for aspects like marketing, storage, labour force requirements and many other elements in the production line. Further and most important is the fact that we have determined a harvesting strategy that does not leave the population vulnerable to extinction. The farmer will continue enjoying long-term revenue from his crop.

1.11 Outline of the thesis

In this chapter, we have introduced the aspect of population modeling, giving a preview on the areas of concern that we will address in later chapters. We also introduce bamboo, its properties, uses and cultivation process. Further, we state our objectives and justify them.

In Chapter Two, we discuss useful literature that introduces what has been done so far in this area of mathematical population modeling. We have highlighted deficiencies of available work in addressing the problem at hand.

In Chapter Three, we give the methods used in both modeling and bamboo data analysis. Brief theory on the logistic model, stochastic processes, the Itô's Lemma and stochastic calculus are given. This gives way to Chapter Four where we undertake the modeling and analysis of model equations. We also simulate trajectories at various instances of modeling.

In chapter Five we analyze bamboo data and give various comparisons of growth parameters. We proceed to fit the suggested harvesting models and get useful estimates of mean sustainable yield (MSY). Finally in Chapter Six, we summarize our results in line with the stated objectives, and give suggestions for further work.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

One of the goals of modern population ecology is to understand population phenomena in terms of the behavior of individuals. The development of a stand of plants results from the growth of individuals, as altered by interactions among these individuals. We know that plant growth is sigmoid [46, 94], and several sigmoid growth models with biologically interpretable parameters have been proposed to describe the growth of individual plants. These include the logistic and Gompertz models. In most sigmoid growth models, initial growth is exponential, and a negative term reduces the relative growth rate as size increases, resulting in an asymptotic maximum size [101]. The main difference among different sigmoid growth models is the inflection point, the size at which a plant experiences its maximum absolute growth rate [82].

In the Richards model, this inflection point is not fixed but modeled by an additional parameter [75]. This makes the Richards growth model highly flexible and inclusive of most other sigmoid growth models. The Richards growth model generally fits plant growth data well and has been widely used in plant ecology and forestry. Estimating the inflection point as a free parameter seems biologically reasonable since there is no general theory that predicts at what growth stage plants experience their maximum growth rate, and the inflection point has been shown to depend on density [32]. One biologically undesirable feature of the Richards growth model is that when maximum growth rate is obtained early in plant growth, i.e. the growth curve has a low inflection point, initial growth is not exponential [18]. To address this problem, Birch [18], proposed a growth model that has both initial exponential growth and an inflection point parameter. The degree of mathematical flexibility of the Birch growth model is comparable to that of the Richards model and both models have the same number of parameters, so it is expected that the Birch growth model will fit plant growth data as well as the Richards growth model. Furthermore, there are some problems with parameter estimation in the Richards

growth model [18, 73, 92, 101], but since the Birch model does not include a power function of the current size relative to the maximum size, it is expected to have better estimation properties than the Richards model [18].

In a previous study [32], the Richards growth model was used to analyze competition among individual plants by coupling individual growth models such that the saturation term is based on the cumulative size of the whole population and including effects of size-asymmetric competition [98]. Genetic differences in microhabitat and variation in local competition may cause variation in the shape of individual growth curves. Accounting for such variation among individuals is important to reduce biased and auto correlated residuals, which may influence inferences made concerning the parameters of interest [24].

There is increasing interest in including the variation among individuals in the modeling of population growth with hierarchical models, in which some parameters are estimated at the level of the population, whereas other parameters are assumed to vary among individuals according to a specific distribution (e.g. a normal distribution), where the mean and the variance of the distribution can be estimated [26, 38, 81].

Diameter at breast height (dbh) and crown width are important tree characteristics and an accurate prediction of tree dimensions has become prominent as analysis techniques, models, and other statistical tools to allow for the rapid evaluation of extensive volumes of data. Some variables (e.g., diameter or age) are easy to measure with simple instruments and it is widely used by forest inventories. However, a number of studies have shown that other variables which are not so easily obtained are also good predictors of forest dynamics and they can improve the reliability of tools like growth and yield models. One of these variables is crown size, which has received increasing attention as a means to estimate tree growth [20]. Crown diameter prediction has been studied by many researchers [12, 13, 15] as a measure of plant growth. Curve fitting in an effort to determine appropriate growth distributions of tree growth parameters is another interesting area that has been addressed [89].

Measurement of tree crown width is more difficult and more time consuming than that of dbh [10]. Crown width is used in tree and crown level growth-modeling systems, where simple competition indices are not available to adequately predict recovery from competition when a competitor is removed [94]. Crown width is also used in calculating competition indices based on crown overlap [17, 33] and predicting above ground biomass. Crown width models can be formulated from open-grown trees or from stand-grown trees [40]. Equations for predicting the dimensions of crowns in open locations consider maximum biological potential, and so are known as "maximum crown width" (MCW) equations, while those for stand-grown trees which generally have a smaller crown

due to competition, are called "largest crown width" (LCW) equations [44]. LCW models predict the actual size of tree crowns in forest stands, and have many applications including estimations of crown surface area and volume in order to asses forest health [96], tree- crown profiles and canopy architecture [45, 57], forest canopy cover [39] and the aboveground biomass. Modeling crown diameter as a simple linear model between crown width and diameter at breast height is often adequate [23, 66, 87].

The development of equations to predict dbh, height, crown diameter, crown height, leaf area and other measures will enable researchers and forest managers to model cost and benefit, analyze alternative management scenarios and determine the best management practices for substancial forestry [58]. Peper et al [68] suggests a logarithmic regression equation for predicting growth measures using age or diameter at breast height (dbh).

Addison et al [1] evaluated the utility of using variable-top, total and merchantable tree green weight equations for predicting crown biomass on operational harvests in southern pine stands. Stands at three locations in Georgia were inventoried prior to an operational thinning or clear cut. Round-wood and biomass amounts were estimated using the inventory data with appropriate weight equations. Actual tonnage of round-wood and biomass were monitored by product during the operational harvest. A post-harvest inventory was then performed to assess the residual stands after thinning as well as any residual material left on the ground at each site. Pre-harvest estimates were compared to the sum of material harvested and remaining on site to assess the utility of this approach to accurately estimate biomass available for harvest in an operational setting.

2.2   Growth and harvest modeling

Brauer and Sanchez [21], [22] investigated the effect of constant rate harvesting periodic harvesting in periodic environments on the asymptotic behaviour of several continuous population models. In both cases, a great deal of attention was given to the continuous logistic model.

The maximum sustainable yield (MSY) for the continuous logistic model under different harvesting strategies has been widely investigated in the literature [14, 25, 37, 101]. Zhang et al. [102] argued that continuous harvesting under the logistic model is superior when compared with an impulsive harvesting.

Alvares and Shepp [4] investigated optimal harvesting for stochastically fluctuating populations. This was done under unbounded and bounded harvesting rates. They recommended partial harvesting as opposed to clean cuts.

Liu [54] considered a stochastic logistic model with impulsive perturbation. He added an equation to the known Stochastic Logistic Equation (SLE) which checks fluctuations in population size at some future time t^+ compared to current time t. Properties of solutions were also discussed.

Pasquali [65] studied a variant of the stochastic logistic equation where she determined the stability of stationary solutions. She simulated paths of trajectories at various levels of noise and noted that at large σ, the system goes extinct.

Zhang et al [102] established a model for optimal impulsive harvesting policy for an autonomous population using the logistic equation. They added a Dirac impulsive function to a proportional harvest term.

Soboleva and Pleasants [88] used non-linear stochastic differential equations for growth process of biological populations. In particular they studied the Gompertz and logistic stochastic models. They showed that initial conditions have more effect on variable than the mean process.

Shah [83] studied a stochastic logistic model for fish growth. His model included an oscillating carrying capacity. He however stated that in the long run, the system will follow an approximately stable path. Qi and Mao [69] developed a stochastic population model where they showed that noise suppresses explosion in population dynamics. They also showed that the model is stochastically bounded.

Allen [3] discussed continuous time SIR malaria nodes under Markov chain, deterministic, and stochastic procedures. Simulated sample paths were obtained under the three conditions. It was noted that while there was variance in trajectories under the three systems, the general trend was the same.

Hritonenko et al. [62] discussed a bang-bang regime in optimal harvesting of size-structured populations. The conclusion was that a linear age structural population model is similar in optimal population trajectory to the non-linear size structured model. It was recommended selective logging between two common logging regimes; clear cutting and selective logging.

Li and Wang [53] considered some optimal harvesting policy for a general stochastic logistic population model. The optimization was computed for the objectives of maximum retained profit. Various powers of the density term were considered and their effect on the behavior of the system discussed.

Liu and Wang [54] studied stationary distribution, ergodicity and extinction of a generalized stochastic logistic system. They noted that when the intensity of white noise

is sufficiently large, the system is extinctive. However, for sufficiently minor intensities, there is a stationary and ergodic population distribution.

Zou et al [99] discussed the Ergodic method of optimal harvesting for a stochastic Gompertz type diffusion process. They also showed equivalency of results on optimal harvesting strategy by mean compared to use of the ergodic method.

A stochastic Markov chain model for bamboo harvest forecasting was also reported by [9]. In their model, transition probabilities derived from bamboo growth data were used to construct a Markov chain. This was analyzed and steady state probabilities of the system derived. These were in turn used to predict future values of harvest. The shortfall in the stochastic model is that it used data only within three years of growth. It therefore was able to determine only the size of the first harvest. The researchers suggested further research on what happens after the first harvest.

Doust and Saraj [35] considered harvesting a logistic population with constant growth and harvesting rate h. They further explored a case of variable harvest Y dependent on the population x such that $Y = hx/(x + 1)$. They assumed that the carrying capacity $K = 1$. Further, a theorem describing the stability of a real model of single species was proved.

Jingliang and Wang [49] studied almost sure permanence of stochastic single species models. They included some numerical simulations to support their main results.

Morteza and Kiaee [60] studied the stochastic dynamical theta-logistic population model with application to Iranian population data between 1921 to 2011. Some preliminaries of stochastic calculus were presented. They also simulated growth paths for various values of theta.

Recently, Anderson et al. [8] suggested that the carrying capacity of an environment is stochastic. They used SDEs to determine the paths, variances of solutions and distributions using the Monte-Carlo method. They however did not solve the equation to give explicit solutions. They claimed that solutions of some variants of stochastic logistic equations had not been found.

Most recently, Yoshioka [100] suggested a simplified stochastic optimization model for logistic dynamics with control dependent carrying capacity. This was motivated by the algae population management problem. It was noted that a given amount of flow must reach some threshold for it to be able to break off and push away algae at a river bed.

2.3 Agricultural Production Systems Simulator (AP-SIM)

APSIM [11] is a cropping systems modeling environment specially designed to allow a plug-in-pull-out approach for the integration of various simulation models via a common modeling protocol [59]. It is a product of the Agricultural Production Systems Research Unit (APSRU). APSIM can be configured with modules suitable for the simulation of many different systems. Whilst these initially concentrated upon dry-land cropping systems, APSIM's usage has broadened and now it is also being used in the study of forestry [67], agro forestry [47] and pasture [86] systems. Horticultural crop models are now being included into the suite of crop models available within APSIM. APSIM-Broccoli is one such model.

APSIM software is a modular modeling framework that has been developed by APSRU (Agricultural Production Systems Research Unit) in Australia. APSIM was developed to simulate biophysical processes in farming systems, particularly as it relates to the economic and ecological outcomes of management practices in the face of climate risk.

APSIM is structured around plant, soil and management modules. These modules include a diverse range of crops, pastures and trees, soil processes including water balance, N and P transformations, soil pH, erosion and a full range of management controls. APSIM resulted from a need for tools that provided accurate predictions of crop production in relation to climate, genotype, and soil management factor while addressing the long-term resource management issues.

2.4 Conclusion

Much of work has been reviewed in this section. No doubt, there is a lot of mathematical research that has been done with impressive results documented. However, other than some applications in the fishing industry and human population projections, much of it remains mathematical proficiency. In all the literature, a good amount of theoretical effort has been made to derive various population dynamic equations. Plant population dynamics using stochastic differential equations is rare in literature. From the literature cited, it was found that different variants of the logistic model have been studied. There is however no clear explanation on which particular format fits a given natural phenomenon. It is therefore left to the imagination of the reader to decide where and when to use a particular formulation.

In our study, we attempt to model as we fit the system to real data and test how well the data agrees with the model dynamics. Different populations behave differently, so a model that fits one instance of a population does not necessarily fit another. Furthermore, tailor made models that have minimum assumptions are more befitting that general ones. Specifically, we construct bamboo growth and harvest models where we drop the assumption of constant growth rate, constant harvest rate and deterministic solutions. We use stochastic differential equations, incorporate harvest and study the systems behavior, steady state (equilibrium) points and extinction in finite time.

Further, we note that all harvest models we were able to sample in literature assume that the initial population is large enough so that harvesting starts immediately. From our observation with bamboo, where one seedling is planted, from which tens of culms emerge after some time, harvesting cannot start immediately. In fact, a harvest free period of 3 years must be accorded in order for the harvest to be mature and useful. In our work, we therefore suggest time delayed harvesting models, where the first three years are pure growth phase and for the rest of the years, the harvest model picks up.

CHAPTER THREE

METHODOLOGY

3.1   Introduction

In this chapter, the methods used during the study are introduced. Stochastic processes, Ordinary differential equations and stochastic differential equations are defined. Methods of solving SDEs are also given. Furthermore, various data analysis methods are also discussed.

3.2   Essential Theory

3.2.1   Random Walk

Let $\{X_i; i \geq 1\}$ be a sequence of of IID random variables, and let $S_n = X_1 + \cdots + X_n$. The integer-time stochastic process $\{S_n; n \geq 1\}$ is called a random walk, or, more precisely, the one dimensional random walk based on $\{X_i; i \geq 1\}$.

For a given n, S_n is simply a sum of IID random variables, but here the behavior of the entire random walk process, $\{S_n; n \geq 1\}$, is of interest .Thus, for a given real number $\alpha > 0$, we might want to find the probability that the sequence $\{S_n; n \geq 1\}$ contains any term for which $S_n \geq \alpha$ (that is the threshold at α is crossed) or to find the distribution of the smallest n for which $S_n \geq \alpha$.

A simple random walk

Suppose $X_1, X_2, \ldots$ are IID binary random variables, each taking on the value 1 with probability p and -1 with probability $q = 1 - p$. Letting $S_n = X_1 + \cdots X_n$, the sequence

of sums $\{S_n; n \geq 1\}$, is called a simple random walk. Note that S_n is the difference between positive and negative occurrences in the first n trials, and thus a simple random walk is little more than a notational variation on a Bernoulli process. For the Bernoulli process, X takes the values 1 and 0, whereas for a simple random walk X takes on the values 1 and -1. For the random walk, if $X_m = 1$ for m trials, then $S_n = 2m - n$, and

$$Pr(S_n = 2m - n) = \binom{n}{m} p^m q^{n-m} \tag{3.1}$$

3.2.2 Stochastic Processes

Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, and that $X : \Omega \to \mathbb{R}$ is a random variable. Recall that this means that Ω is a space, $\mathcal{F}$ is a σ-algebra of subsets of Ω, P is a countably additive, non-negative measure on $(\Omega, \mathcal{F})$ with total mass $P(\Omega) = 1$, and X is a measurable function, i.e.,

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

for every Borel set $B \in B(\mathbb{R})$.

A stochastic process is simply a collection of random variables indexed by time. It will be useful to consider separately the cases of discrete time and continuous time. A discrete time stochastic process $X = \{X_n, n = 0, 1, \ldots\}$ is a countable collection of random variables indexed by the non-negative integers, and a continuous time stochastic process $X = \{X_t, 0 \leq t < \infty\}$ is an uncountable collection of random variables indexed by the non-negative real numbers.

3.2.3 Martingales

A martingale is defined as an integer-time stochastic process $\{X_n; n \geq 1\}$ with properties that $\mathbb{E}[|X_n|] < \infty$ for all $n \geq 2$ and

$$\mathbb{E}[X_n | X_{n-1}, X_{n-2}, \ldots, X_1] = X_{n-1} \tag{3.2}$$

for all $n \geq 2$.

The name martingale comes from gambling terminology where martingales refer to gambling strategies in which the amount to be bet is determined by the past history of winning or losing. If one visualizes X_n as representing the gambler's fortune at the end of

the nth play, the definition above means, first, that the game is fair (in the sense that the expected increase in fortune from play $n-1$ to n is 0.), and, second, that the expected fortune on the nth play depends on the past only through the fortune on play $n-1$.

$$\mathbb{E}[X_n|X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_1 = x_1] = x_{n-1} \tag{3.3}$$

for all possible sample values $x_1, x_2, \ldots, x_{n-1}$.

The second is that

$$E[X_n|X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_1 = x_1]$$

is a function of the sample values $x_1, \ldots, x_{n-1}$ and thus $E[X_n|X_{n-1}, \ldots, X_1]$ is a random variable which is a function of the random variables $X_1, \ldots, X_{n-1}$.

One example of of a martingale is a zero mean random walk. Since if $X_n = Y_1 + \cdots + Y_n$, where the Y_i are IID with zero mean, then

$$E[X_n|X_{n-1}, \ldots, X_1] = E[X_n] + X_{n-1} = X_{n-1}.$$

3.2.4 Brownian motion

The conventional way to model the velocity of a particle $V(X(t), t)$ is by using Brownian motion $(B_t)_{t\geq 0}$ with the properties:

(i) $B_t \sim N(0, t)$

(ii) $B_0 = 0$. The process starts at zero.

(iii) $dB_t = B_{(t+dt)} - B_t$ is independent of B_t, that is, B_t is stationary with independent increments.

(iv) The increment $B_t - B_s$ with $0 < s \leq t$ is $N(0, (t-s))$.

(v) $cov[B_t, B_s] = \min\{t, s\}, \forall s, t \in \mathbb{R}^+$

(vi) B_t is a martingale, that is, the expected value of the future is the present.

$$\mathbb{E}[B_{t+s}|B_t] = B_t$$

(vii) B_t is continuous everywhere.

(viii) B_t is fractal

3.2.5 Beta function

The Beta function is given by

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \qquad (3.4)$$

Alternatively

$$B(a,b) = \int_0^\infty \frac{x^{a-1}}{(1+x)^{a+b}}dx \qquad (3.5)$$

with $x > 0, a, b \in \mathbb{R}^+$

3.2.6 Gamma function and the Gamma density

The Gamma function is given by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^x dx \qquad (3.6)$$

with identities

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \text{ and } \Gamma(1) = 1, \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, x > 0, a, b \in \mathbb{R}^+$$

A continuous random variable X has a Gamma distribution if

$$f(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases} \qquad (3.7)$$

3.2.7 Itô's lemma

Let X be an Itô process with stochastic differential equation

$$dX(t) = \mu(t)dt + \sigma(t)dW(t) \qquad (3.8)$$

Further assume that we are given a $C^{1,2}$ function $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ Define a new process Z by

$$Z(t) = f(t, X(t)) \qquad (3.9)$$

Then Z has a stochastic differential equation given by

$$
\begin{aligned}
df(t, X(t)) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX(t) + \frac{\partial^2 f}{2 \partial x^2}[dX(t)]^2 \\
&= \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{\sigma^2 \partial^2 f}{2 \partial x^2} \right) dt + \sigma \frac{\partial f}{\partial x} dW(t)
\end{aligned}
\tag{3.10}
$$

where the term $\mu \frac{\partial f}{\partial x}$ is shorthand for $\mu(t) \frac{\partial f}{\partial x}(t, X(t))$ and so on. Note that formally

$$
\begin{aligned}
[dX(t)]^2 &= [\mu dt + \sigma dW(t)]^2 \\
&= \mu^2[dt]^2 + 2\mu\sigma[dt][dW(t)] + \sigma^2[dW(t)]^2 = \sigma^2 dt
\end{aligned}
\tag{3.11}
$$

where we used the following multiplication table

$\times$	dt	$dW(t)$
dt	0	0
$dW(t)$	0	dt

In the special case where the function $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable we get:

$$
df(X(t)) = \left(\mu f'(X(t)) + \frac{1}{2}\sigma^2 f''X(t) \right) dt + \sigma f'X(t)dW(t)
\tag{3.12}
$$

Consider the population growth model

$$
\frac{dN_t}{dt} = a_t N_t
\tag{3.13}
$$

where N_0 is given. We choose $a_t = r_t + \sigma \frac{dW_t}{dt}$ in order to include uncertainty in growth rate in the model. Let us assume that $r_t = r$ we include the uncertainty in the model. constant. By the Itô interpretation, equation (3.13) is equivalent to

$$
dN_t = rN_t dt + \sigma N_t dW_t
\tag{3.14}
$$

Equivalently

$$
\frac{dN_t}{N_t} = rdt + \sigma dW_t
\tag{3.15}
$$

$$
\int_0^t \frac{dN_s}{ds} = rt + \sigma W_t
\tag{3.16}
$$

where $W_0 = 0$. The evaluation of the integral on the left hand side requires the use of the Itô formula for the function

$$
g(t, x) = \ln x, \quad x > 0
$$

23

In this case get

$$d(\ln N_t) = \frac{1}{N_t}dN_t + \frac{1}{2}\left(-\frac{1}{N_t^2}\right)(dN_t)^2$$
$$= \frac{dN_t}{dt} - \frac{1}{2N_t^2}\sigma^2 N_t^2 dt$$
$$= \frac{dN_t}{dt} - \frac{1}{2}\sigma^2 dt$$

Therefore

$$\frac{dN_t}{dt} = d(\ln N_t) + \frac{1}{2}\sigma^2 dt \qquad (3.17)$$

Equating equation (3.17) to equation (3.15) we find that

$$d(\ln N_t) + \frac{1}{2}\sigma^2 dt = rdt + \sigma dW_t$$

It follows that

$$d(\ln N_t) = \left(r - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t \qquad (3.18)$$

$$\int_0^t d(\ln N_s) = \int_0^t \left[\left(r - \frac{1}{2}\sigma^2\right)ds + \sigma dW_t\right]$$

$$\ln\left(\frac{N_t}{N_0}\right) = \left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t.$$

Hence

$$N_t = N_0 e^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t} \qquad (3.19)$$

3.2.8 Fokker-Planck equation

The Fokker-Planck equation is a partial differential equation that describes the time evolution of the probability density function $P(X_t, t)$ of the velocity of a particle under the influence of drag forces $f(X_t, t)$ and random forces $g(X_t, t)$. The velocity can also be seen as a rate of growth of some population that can be represented as a stochastic Itô process

$$dX_t = f(X_t, t)dt + g(X_t, \sigma)dW \qquad (3.20)$$

For a uni-dimensional process, $P(X_t, t)$ is given by the partial differential equation

$$\frac{dP(X_t, t)}{dt} = -\frac{\partial f}{\partial x}P(X_t, t) + \frac{\partial^2 D}{2\partial x^2}P(X_t, t) \qquad (3.21)$$

where $D = g^2(X_t, t)$. It is also known as the Kolmogorov forward equation.

The Ornstein-Uhlenbeck process is a process defined as

$$dX_t = -\mu X_t dt + \sigma dW_t \tag{3.22}$$

Its associated Fokker-Planck equation is

$$\frac{dP(X_t, t)}{dt} = \mu \frac{\partial}{\partial x}[X_t P(x, t)] + \frac{\sigma^2 \partial^2}{2 \partial x^2} P(x, t) \tag{3.23}$$

We can use Itô's theorem to transform equation (3.22).

Let $y_t = X_t e^{\mu t}$ then

$$X_t = y_t e^{-\mu t}, \quad \frac{\partial y_t}{\partial t} = \mu X_t e^{\mu t}, \quad \frac{\partial y_t}{\partial X} = e^{\mu t} \text{ and }, \quad \frac{\partial^2 y_t}{\partial X^2} = 0$$

Substituting this into Itô's Lemma, we get

$$dy_t = \sigma e^{\mu t} dW_t \tag{3.24}$$

Integrating on both sides of equation (3.24) we get

$$y_t = y_0 + \sigma \int_0^t e^{\mu s} dW_s$$

Reverting to initial notation we have

$$X_t = X_0 e^{-\mu t} + \sigma e^{-\mu t} \int_0^t e^{\mu s} dW_s \tag{3.25}$$

To characterize X_t we need its mean and variance. Thus we compute

$$\mathbb{E}[X_t] = X_0 e^{-\mu t} + \sigma e^{-\mu t} \int_0^t e^{\mu s} \mathbb{E}[dW_s] \tag{3.26}$$

so

$$\mathbb{E}[X_t] = X_0 e^{-\mu t} \tag{3.27}$$

Further, given that $Var[X_t] = \mathbb{E}[X_t^2] - (\mathbb{E}[X_t])^2$ we have

$$Var[X_t] = \frac{\sigma^2}{2\mu}(1 - e^{-2\mu t}) \tag{3.28}$$

Thus the distribution of X_t is Gaussian with the derived mean and variance. Explicitly,

$$P(X_t, t) = \frac{1}{\sqrt{\frac{\pi \sigma^2}{\mu}(1 - e^{-2\mu t})}} \exp\left[\frac{-\mu(X_t - X_0 e^{-\mu t})^2}{\sigma^2(1 - e^{-2\mu t})}\right] \tag{3.29}$$

is the density function of the Ornstein Uhlenbeck process.

3.2.9 Stochastic population models

Let the quantity of a measure at time t be X_t. Consider a small subsequent time interval dt during which X_t changes to $X_t + dX_t$, then the intrinsic growth rate is

$$\frac{dX_t}{dt} = R(X_t, t) \tag{3.30}$$

If the rate of change is deterministic then

$$dX_t = R(X_t, t)dt \tag{3.31}$$

However, the rate of change is generally not deterministic. It often is affected by many factors and uncertainties including environmental disturbances. These are stochastic in nature. We model uncertainty by adding a stochastic part as follows:

$$dX_t = R(X_t), t)dt + V(X_t, t)N(0, dt) \tag{3.32}$$

Here the random change $V(X_t, t)$ is the stochastic part modeled as a normal distribution with mean 0 and variance V^2.

The stochastic model can therefore be written as

$$dX_t = R(X_t), t)dt + V(X_t), t)dB_t \tag{3.33}$$

which can be rewritten as

$$dX_t = RX_t dt + X_t dB_t \tag{3.34}$$

This is a stochastic differential equation. Here R and V can be constants or linear functions in X_t and t. That is $R = b + aX_t$ and $V = \sigma X_t$. In this case, the equation becomes

$$dX_t = X_t([b + aX_t]dt + \sigma dB_t) \tag{3.35}$$

Consider the exponential growth differential model

$$dX_t = \mu X_t dt \tag{3.36}$$

It is written as a stochastic differential equation (SDE) as

$$dX_t = \mu X_t dt + \sigma X_t dB_t \tag{3.37}$$

The SDE has an explicit solution (see example in section 3.2.8)

$$X_t = X_0 e^{(\mu - 0.5\sigma^2)t + \sigma B_t} \tag{3.38}$$

We note that if $\sigma = 0$ we have the simple exponential solution

$$X_t = X_0 e^{\mu t}. \tag{3.39}$$

Now for solution of the stochastic exponential equation, when

(i) $\mu > 0.5\sigma^2$, the population $X_t \to \infty$ exponentially with probability 1.

(ii) $\mu < 0.5\sigma^2$, the population $X_t \to 0$ exponentially with probability 1.

(iii) $\mu = 0.5\sigma^2$, then $\liminf\limits_{t\to\infty} X_t = \infty$ while $\limsup\limits_{t\to\infty} X_t = 0$ with probability 1.

We note that in the ODE model $\mu > 0$ leads to population explosion while in the SDE the population goes extinct almost surely. This illustrates the importance of considering stochastic modeling over deterministic models.

3.3 Bamboo data

The study was carried out in four sites in the former South Nyanza region of Kenya. These were Rangwe and Suba (in Homa Bay County), Ngege and Kuria/Ikierege (in Migori County). To conduct this study, 120 farmers in the region were selected forming four bamboo farmers' cooperative societies each with 30 farmers. The farmers were selected on the basis of whether or not one was a tobacco farmer, gender, age, poverty status, farming scale, access to water and the willingness to provide land for the bamboo experimentation. A total of 2,420 bamboo cuttings consisting of 1210 giant and 1210 *bambusa* species were planted. Each farmer was given 20 bamboo cuttings (that is, 10 each of giant bamboo and 10 of common bamboo). Each cutting was planted in a cubic hole measuring $0.6 \times 0.6 \times 0.6$ m. Spacing between two-bamboo clumps was 5 m. 50% of bamboo clumps were randomly selected and tagged with codes indicating the site, farmer's number, species and the clump number for easy identification and monitoring. Survival rates, number of culms, culm heights and diameters, among other parameters were measured.

3.3.1 Survival proportion

The survival fraction was calculated by counting the number of surviving clumps in every farm after six months of planting the bamboo cuttings. Another rate was calculated after a year of planting. Subsequent rates remained the same since clumps of over one year are already hardy.

3.3.2 Number of culms per clump

A culm is a bamboo pole while a clump is the cluster of poles that emerge from one seedling. The number of culms in the tagged clumps were counted and recorded. This was done at an interval of three months starting from the planting date to ascertain the rate of culm development. After one year, the measures were taken in intervals of six months.

3.3.3 Culm height

Three dominant culms were selected from each of the five selected (tagged) clumps and their heights measured. This was accomplished using a, tape measure for short culms or a Suunto height meter or a straight pole marked in meters for tall ones. Bent stems were straightened so that the actual length of the stem could be measured.

3.3.4 Culm diameter

Twelve months after planting bamboo cuttings of the two species, the diameter of each of the selected bamboo culms was measured using a diameter tape. Progressive measurement was done on new thicker culms.

3.4 Data Collection

On each farm, there were 10 *bambusa vulgaris* and 10 *dendrocalamus giganteus* clumps planted. Five clumps of each species were tagged for consistent data recording. Observations were made on various clumps of each species and the number of culms and new shoots per clump counted at various intervals of time (3 to 6 months). The observations

of the clumps involved visually identifying small, medium, and large culms and measuring the diameters of selected culms using a measuring tape. Based on the measured samples of various sizes of culms of each species, an average culm diameter was calculated for each species.

The average height of the culms of each species was determined in many cases, roughly by observation, as well as accurately, by using a height meter. This method was applied by sampling 3 of the tallest culms per clump, for tagged clumps of each species.

The estimated potential (the number of averagely sized culms per hectare per year) was made with the assumption that, all mature culms per clump are annually harvested. Note that only 3 year old mature culms are harvested and that new culms (year 0, shoots), as well as 1 and 2 year old culms are left to mature.

3.5   Monte-Carlo simulation

Monte Carlo is also suitable for solving complex stochastic problems because it can deal with a large number of random variables, various distribution types, and highly non-linear stochastic models. Different from a physical experiment, Monte Carlo simulation performs random sampling and conducts a large number of experiments on computer. Then the statistical characteristics of the experiments (model outputs) are observed, and conclusions on the model outputs are drawn based on the statistical experiments. In each experiment, the possible values of the input random variables are sampled (generated) according to their distributions. Then the values of the output variable are calculated through the performance of function transformation of samples of input random variables. With a number of experiments carried out in this manner, a set of samples of output variable are available for the statistical analysis, which estimates the characteristics of the output variable.

A Monte Carlo method consists of representing the solution of a problem as a parameter of a hypothetical population, and using a random sequence of numbers to construct a sample of the population, from which statistical estimates of the parameter can be obtained. This is also sometimes referred to as stochastic simulation.

The objective is to estimate an integral $\int_x f(x)dx$ which is analytically intractable. In practice, using Monte Carlo simulation to investigate the properties of an estimator or test follows the procedure:

(i) Generate a sample (size N) using a population model (the data generating process).

29

(ii) Calculate your estimates/statistics.

(iii) Repeat this M times (when you are interested in the tails of a distribution you need more replications than when you are interested in the center).

Monte Carlo simulation can be considered as a methodical way of doing a what-if analysis. In Monte Carlo simulation, we identify a statistical distribution which we can use as the source for each of the input parameters. Then, we draw random samples from each distribution, which then represent the values of the input variable. The steps performed for the Monte Carlo simulation of a physical process [60, 63] are further explained.

3.6 Random number generation

After identifying the underlying distributions for the input parameters of a simulation model, we generate random numbers from these distributions. The generated random numbers represent specific values of the variable. Next, we will discuss the most common method for generating random numbers from discrete and continuous distributions.

3.6.1 Generating random numbers from a distribution function

a) Inverse Transformation Method

This provides the most direct route for generating a random sample from a distribution. In this method, we use the inverse of the cummulative density function (CDF) (for continuous distributions) or cummulative mass function (CMF) (for discrete distributions), and convert a random number between 0 and 1 to a random value for the input distribution. The process can be Mathematically described as follows:

Let X be a continuous random variable (which we want to generate) following a PDF f. Let the cumulative probability distribution function (CDF) for the variable be denoted by F which is continuous and strictly increasing in $(0, 1)$. Let F^{-1} denote the inverse of the function F, which is often called inverse CDF. Then, the following two steps will generate a random number X from the CDF f.

(i) Generate $U \sim U(0, 1)$

(ii) Return $X = F^{-1}(U)$.

Since $0 \leq U \leq 1, F^{-1}(U)$ always exists.

b) Generating a random sample from a dataset: Bootstrapped Monte Carlo

Often it is not possible to obtain an underlying distribution for an input variable in a simulation model. This can be because of the complicated shape of the original distribution (like non-convex or multi-modal), scarcity of data (for example, destructive testing or costly data) and so on. In those cases, we might end up with nothing more than a few historical values for the input parameter. Here bootstrapped Monte Carlo simulation (often called bootstrapping) can be used to generate a random sample. In bootstrapping, we do not really generate new random numbers. Instead, we repeatedly sample the original dataset to choose one of the data points from the set (choose a number with replacement). For many datasets, this method provides good result for simulation purposes. For bootstrapped Monte Carlo simulation, one has to still use a uniform random number generator (RNG) to generate integer random numbers among the indices of an array, which is being used for storing the original dataset.

3.6.2   Monte Carlo simulation output analysis

The result of the Monte Carlo simulation of a model is typically subjected to statistical analysis. As mentioned before, for each set of random numbers (or trials) generated for each of the random variable, we use the model formula to arrive at a trial value for the output variable(s). When the trials are complete, the stored values are analyzed [81]. Averaging trial output values result in an expected value of each of the output variables. Aggregating the output values into groups by size and displaying the values as a frequency histogram provides the approximate shape of the probability density function of an output variable. The output values can themselves be used as an empirical distribution, thereby calculating the percentiles and other statistics. Alternatively, the output values can be fitted to a probability distribution, and the theoretical statistics of the distribution can be calculated. These statistics can then be used for developing confidence bands. The precision of the expected value of the variable and the distribution shape approximations improve as the number of simulation trials increases.

All the operations discussed above have been incorporated in statistical software and can be readily applied without programming them first. In R, there are various packages for the operations discussed [84, 86].

3.7   Data analysis

Collected time series data were entered into spreadsheets and cleaned using Excel. Data analysis was done in R and SPSS. Consistency of collected data was affected by death of some clumps and loss of tags. Various statistical tools were used during data analysis. First, a descriptive analysis of growth measures was performed producing frequency tables and bar charts. Then time series data were fitted with regression models (curve fitting). Further, the data collected were used in determining probability distribution parameters for use in simulation. Comparison of performance of growth measures in different study sites and between the two species was accomplished by correlation analysis, and tests of hypotheses.

CHAPTER FOUR

BAMBOO POPULATION GROWTH AND HARVEST MODELING

4.1 Introduction

Starting from the classical Verhulst ordinary differential equation the harvesting component was introduced and the model characteristics at equilibrium determined. An extension to the stochastic differential equations realm was explored with explicit solutions computed at equilibrium. The models were constructed starting with the following general assumptions.

   i. That shooting is continuous and uniform and that the shooting rate (which we shall refer to as intrinsic population growth rate) will be constant throughout the life of the clump.

   ii. That a culm that is 3 years old is ready for harvesting. Thus we assume harvesting occurs for oldest culms in a clump before new ones are harvested. So harvesting is discriminate on age.

   iii. Harvesting is continuous just like growth.

   iv. The rate of harvesting is constant.

   v. There is no competition between bamboo and other plants in the farm.

   vi. The carrying capacity of a clump is constant.

4.2 The logistic model

In 1838 the Belgian mathematician Verhulst introduced the logistic equation, which is a kind of generalization of the equation for exponential growth but with a maximum value

for the population. Pierre Francois Verhulst [95] published the equation,

$$\frac{dN_t}{dt} = rN_t \left[1 - \frac{N_t}{K}\right] \tag{4.1}$$

where N_t is the number of individuals at time t, r the intrinsic growth rate and K is the maximum number of individuals that the environment can support. The solution to this logistic equation is called the logistic equation.

$$N_t = \frac{N_0 K}{N_0 + (K - N_0)e^{-rt}} \tag{4.2}$$

Now suppose the growth rate r is stochastic, with population size N_t coupled with noise dB_t then the SDE (4.3) best describes the evolution of the system.

$$dN_t = rN_t \left[1 - \frac{N_t}{K}\right] dt + \sigma N_t dB_t \tag{4.3}$$

4.3   The Beverton-Holt logistic model

A famous discrete time model for logistic growth was first proposed by Beverton and Holt [16] for the dynamics of exploited fish populations. It gives the population dynamics as:

$$N_{t+1} = \frac{\lambda K N_t}{K + (\lambda - 1)N_t}, \quad N_0 > 0, \lambda > 1, \text{ and } t \in \mathbb{N} \tag{4.4}$$

N_t is the number of individuals at time t. K is carrying capacity and λ the inherent growth rate. Bohner and Warth [19] studied the characteristics of this equation. Al Sharawi and Rhouma [2] studied the Beverton-Holt model with constant, periodic and conditional harvest. For constant harvest they gave

$$N_{t+1} = \frac{\lambda K N_t}{K + (\lambda - 1)N_t} - \bar{h}, \quad \bar{h} > 0 \tag{4.5}$$

here all notations have same meaning as in (4.4) with $\bar{h}$ being the intensity of harvesting.

For some constant $\alpha > 0$, $N_t = \alpha X_t$ and $\alpha = \frac{K}{\lambda - 1}, \bar{h} = \alpha h$ then (4.5) becomes

$$\alpha X_{t+1} = \left(\frac{\lambda \alpha X_t K}{K + (\lambda - 1)\alpha X_t}\right) - \bar{h} = \frac{\lambda \alpha X_t}{1 + \frac{\lambda - 1}{K}\alpha X_t} - \alpha h$$

$$X_{t+1} = \frac{\lambda X_t}{1 + X_t} - h \tag{4.6}$$

They summarized constant harvest with the objective of maintaining indefinite survival of the population as follows:

Let $\overline{X_{th}} = (\sqrt{\lambda} - 1)$ and $h = h_{th} = (\sqrt{\lambda} - 1)^2$

- If $X_0 \geq \overline{X_{th}}$, then $h_{th} = (\sqrt{\lambda} - 1)^2$ is the optimal intensity of harvesting.

- If $X_0 \leq \overline{X_{th}}$, then $h_{th} = \frac{(X_0(a-1-X_0))}{(1+X_0)}$ is the optimal intensity of harvesting.

4.4 A Modified Beverton-Holt logistic model

Model (equation (4.6)) assumes that harvesting starts at time zero. Suppose fingerlings are introduced in a pond, as many as desirable. Even if they can still be harvested the following day, it cannot be the objective of the farmer. They must be fed and let to grow for some time until they are mature enough. A period of pure growth is therefore desirable. The model does not satisfy our assumption (ii). Thus in our work, we suggest a delayed discrete harvesting model:

$$X_{t+1} = \begin{cases} \dfrac{\lambda X_t}{1 + X_t}, & t = 0, 1, 2, \ldots, d-1 \\ \dfrac{\lambda X_t}{1 + X_t} - h_{t-d}, & t \geq d \end{cases} \tag{4.7}$$

where h_t is the intensity of harvest in the tth generation and d the delay. For bamboo $d = 3$ years. Meanwhile due to the constant harvest assumption $h_{(t+p)} = h_t$ for all $p \in \mathbb{N}$ as long as $t \geq d$. We further propose a stochastic discrete logistic model where the growth rate is random. The equation is

$$N_{t+1} = N_t + \tilde{r} N_t \left(1 - \frac{N_t}{K}\right) - h \tag{4.8}$$

with $\tilde{r} = r + U[-\sigma, \sigma]$ with uniform random numbers between $-\sigma$ and σ.

σ being the standard deviation of the growth rate to be determined empirically. A 100 sample average path for the process is given in Figure 4.2(b). The following are some simulation results with parameter levels determined from data.
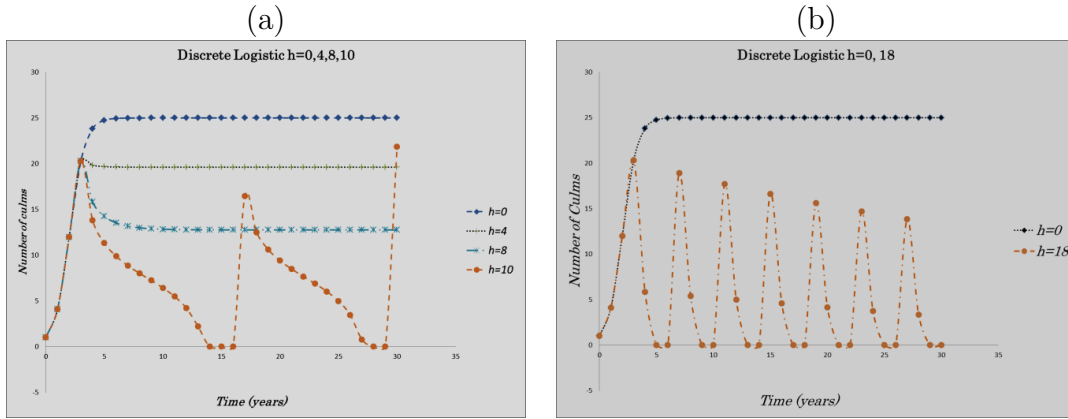
Figure 4.1: Discrete logistic model at various levels of harvest
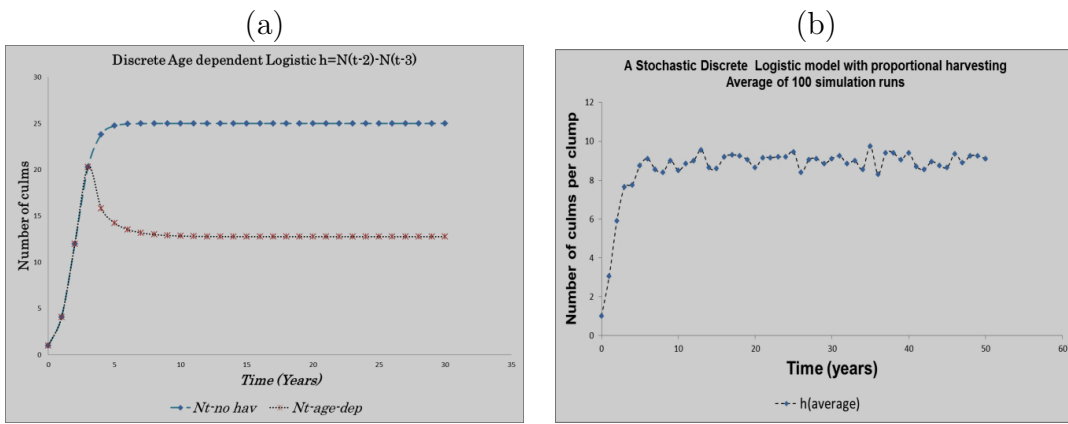
Figure 4.2: Discrete age dependent harvesting logistic model

Figure 4.1(a) shows that the more the harvest, the more unstable the growth system becomes. With no harvest, it takes a logistic growth trajectory. As the number harvested is increased further, the time to a diminishing population is shorter (see Figure 4.1(b)). Clear cutting of the bushes causes spells of no harvest. It is also evident that in the long run the bushes will be extinct. In Figure 4.2(a) the range of upper and lower paths is some constant harvest amount. Figure 4.2(b) gives a simulation of the stochastic discrete logistic distribution with uniform noise $U[-\sigma, \sigma]$ on the growth rate r. It is noted that while noise ensures that the population does not stabilize, there are limits within which the system oscillates. The larger the noise, the wider the volatility band.

4.5 Logistic growth with constant harvest

Next we consider the logistic differential equation with constant harvesting rate H.

$$\frac{dN_t}{dt} = rN_t\left(1 - \frac{N_t}{K}\right) - H \tag{4.9}$$

Where N_t is the population size at time t, K is the carrying capacity, r is the intrinsic population growth rate and H is the harvesting rate (culms per unit time). We find the equilibrium points by letting $\frac{dN_t}{dt} = 0$. That is

$$rN_t - \frac{r}{K}N_t^2 - H = N_t^2 - KN_t + \frac{HK}{r} = 0 \tag{4.10}$$

This has roots

$$l_1 = \frac{K}{2}\left(1 - \sqrt{1 - \frac{4H}{rK}}\right), \ l_2 = \frac{K}{2}\left(1 + \sqrt{1 - \frac{4H}{rK}}\right) \tag{4.11}$$

Consider the value under the square root sign, say

$$\Delta = 1 - \frac{4H}{rK}$$

We study three scenarios

1. When $\Delta = 0$ we have $\frac{4H_C}{rK} = 1$, therefore, $H_c = \frac{rK}{4}$. At this point $E_1 = E_2 = \frac{K}{2}$. This gives the critical harvest per unit time H_c with which the population size remains constant provided $N_0 > E_1$ that is, $N_0 > \frac{K}{2} + \frac{K}{2}\sqrt{1 - \frac{4H}{rK}}$ since we have let $\sqrt{1 - \frac{4H}{rK}} = 0$, then the condition is $N_0 > \frac{K}{2}$. For us, $N_0 = 1$. So there should be no harvest until $N_{t_c} \geq \frac{K}{2}$ that is; for $K = 25$ we wait until a critical $N_{t_c} \geq 13$ culms population is reached. We will compute t_c after finding the expression for N_t.

2. When $\Delta < 0$ the roots are complex .Since we are dealing with a system in $\mathbb{R}^+$ such solution is irrelevant.

3. When $\Delta > 0$, E_1 and E_2 are real valued and distinct. H_m, $H_m < \frac{Kr}{4}$ harvesting in moderation ensures a persistent population.

Next we determine N_t considering the favorable situation: $H < \frac{Kr}{4}$. At equilibrium we have

$$\frac{dN_t}{dt} = \frac{-r}{K}(N_t - E_1)(N_t - E_2) \tag{4.12}$$

Therefore

$$\frac{dN_t}{[N_t - E_1][N_t - E_2]} = \frac{-r}{K}dt \tag{4.13}$$

To make the LHS into partial fractions we have:

$$\frac{1}{[N_t - E_1][N_t - E_2]} = \frac{A(N_t - E_2) + B(N_t - E_1)}{[N_t - E_1][N_t - E_2]} \tag{4.14}$$

We need to find A and B so that we make partial fractions.

$$\frac{[A+B]N_t - [E_2 A + E_1 B]}{[N_t - E_1][N_t - E_2]} = \frac{1}{[N_t - E_1][N_t - E_2]} \qquad (4.15)$$

Since there is no term in N_t on the RHS we have $A + B = 0$ hence $A = -B$ The constant part

$$-(E_2 A + E_1 B) = 1$$

Placing $A = -B$ we have $-(BE_1 - BE_2) = 1$

$$-B(E_1 - E_2) = 1, \quad \Rightarrow B = \frac{-1}{-E_2 + E_1}, \quad \Rightarrow A = \frac{1}{-E_2 + E_1}$$

so we have

$$\frac{1}{E_1 - E_2}\left(\frac{dN_t}{N_t - E_1}\right) - \frac{1}{E_1 - E_2}\left(\frac{dN_t}{N_t - E_2}\right) = \frac{-r}{K}dt$$

Integrating on both sides we have

$$\ln|N_t - E_1| - \ln|N_t - E_2| = \frac{-r}{K}(E_1 - E_2)t + c$$

Therefore

$$\ln\left|\frac{N_t - E_1}{N_t - E_2}\right| = \frac{-r}{K}(E_1 - E_2)t + c$$

where c is derived given the initial condition (at $t = 0$, and $N_t = N_0$) Therefore

$$c = \ln\left|\frac{N_0 - E_1}{N_0 - E_2}\right|$$

Let $\frac{N_0 - E_1}{N_0 - E_2} = A$, then $\frac{N_t - E_1}{N_t - E_2} = Ae^{-r\frac{(E_1 - E_2)t}{K}}$. Hence $N_t - E_1 = (N_t - E_2)Ae^{-r\frac{(E_1 - E_2)t}{K}}$

$$N_t = \frac{E_1 - E_2 Ae^{-r\frac{(E_1 - E_2)t}{K}}}{1 - Ae^{-r\frac{(E_1 - E_2)t}{K}}} \qquad (4.16)$$

$$\lim_{t \to \infty} N_t = \frac{E_1 - E_2 Ae^{-\infty}}{1 - Ae^{-\infty}} = E_1 \qquad (4.17)$$

Here we note that harvesting can cause extinction.

Time to extinction is finite if $N_{ex}(t) = 0$. Using equation (4.16).

$$N_{ex}(t) = \frac{E_1 - E_2 Ae^{-r(\frac{E_1 - E_2)t_{ex}}{K}}}{1 - Ae^{-r\frac{(E_1 - E_2)t_{ex}}{K}}} = 0$$

This implies that

$$t_{ex} = \frac{-K \ln \left| \frac{E_1}{E_2 A} \right|}{r(E_1 - E_2)}$$

Therefore

$$t_{ex} = \frac{1}{r\sqrt{1 - \frac{4H}{rK}}} \ln \left| \frac{E_1(N_0 - E_2)}{E_2(N_0 - E_1)} \right| \qquad (4.18)$$

where $E_i, i = 1, 2$ are as in equation (4.11) and H the constant harvest amount.

From equation (4.18) it can be concluded that the larger the constant harvest the shorter the time to extinction. Figure 4.3 shows the relationship between harvest rate and associated time to extinction.
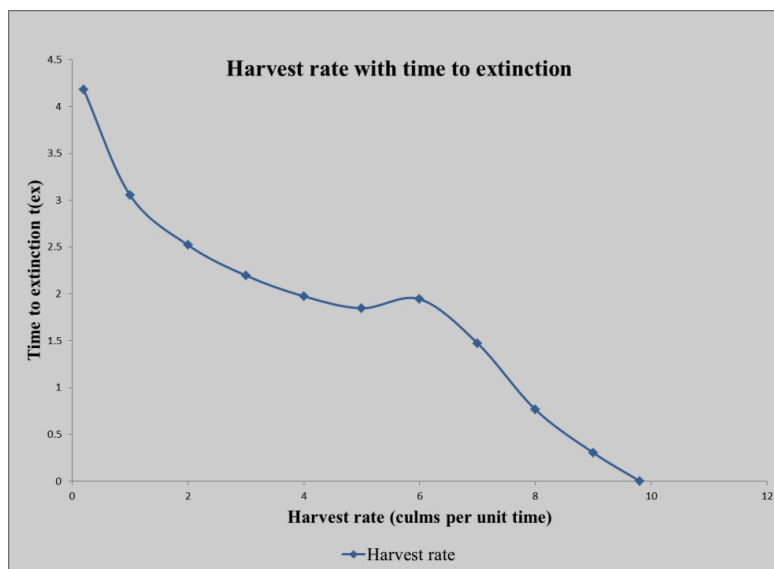
Figure 4.3: Harvest rate effect on time to extinction

Simulation paths of equation (4.16) at various harvest rates are given in Figure 4.4.
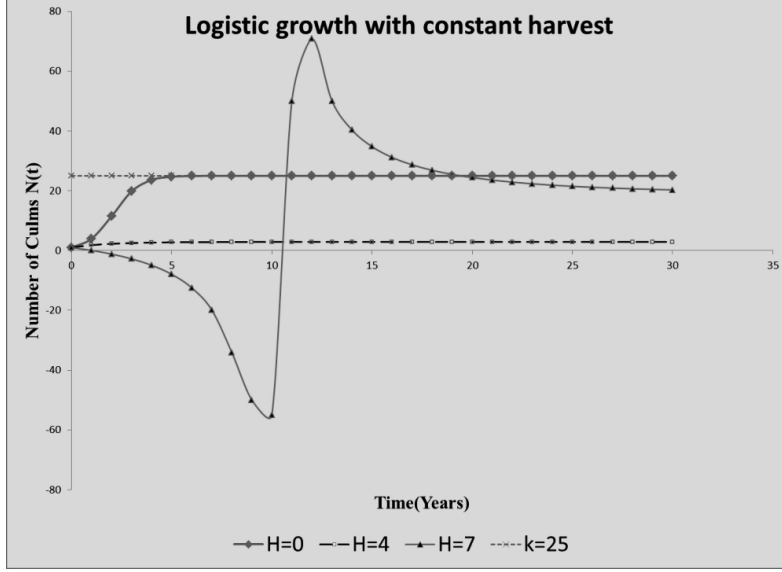
Figure 4.4: Logistic growth with constant harvest

Given equation (4.16), N_t was simulated at $N_0 = 1, r = 1.5686$, and $K = 25$. When there is no harvest, it means $E_1 = E_2 = K/2 = 12.5, A = 1$ equation (4.16) becomes

$$N_t = \frac{KN_0}{N_0 + (K - N_0)e^{-rt}}. \tag{4.19}$$

When there is no harvest, the culm population grows rapidly to a stable maximum $K = 25$. When harvest $H = 4$, the population stagnates at 2 culms throughout and at $H = 7$, the clump goes extinct immediately at the first harvest. This behavior is undesirable. After all, the culms at such age have no significant use since they are immature. The model we suggest basically uses Culm population at t=3 years as N_0 under the harvesting regime so that harvesting can start with some population of mature culms. So the delayed constant harvest equation for bamboo management is

$$N_t = \begin{cases} \dfrac{KN_0}{N_0 + (K - N_0)e^{-rt}}, & 0 \le t < d \\[3mm] \dfrac{E_1 - E_2 A e^{-r\frac{(E_1 - E_2)t}{K}}}{1 - A e^{-r\frac{(E_1 - E_2)t}{K}}}, & t \ge d \end{cases} \tag{4.20}$$

The simulation paths for various values of harvest under this new condition are given in Figure 4.5.
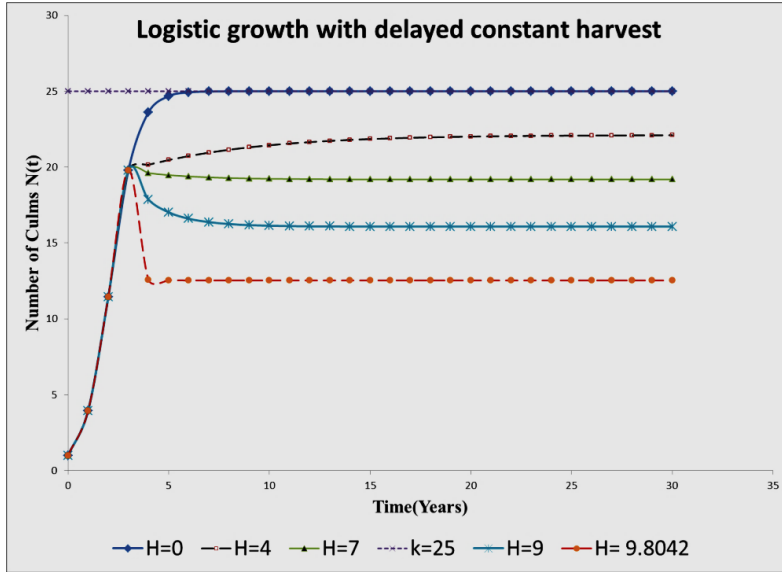
Figure 4.5: Logistic growth with delayed constant harvest

From Figure 4.5, we see that the imposed condition stabilizes the system so that negative populations do not arise. The higher the harvest the closer to 13 culms the stable population gets. The bare maximum average harvest from our simulation is 10 culms per clump. Anything above this number brought about an error. Indeed this is the critical harvest

$$H_c = \left(\frac{rK}{4}\right) = \frac{1.56868(25)}{4} = 9.80425 \approx \text{ culms} \qquad (4.21)$$

At this point, an equilibrium mean population of 13 culms is maintained over the years. An investigation on extinction shows that the system is persistent as long as the harvest is below the critical value. If we harvest 10 culms per year from each clump on average, with the population at the start of harvesting $N_0^* = 19.8 \approx 20$ culms, time to extinction $t_{ex} = -3.46$ years. This is actually the point where the curve crosses the time axis in Figure 4.5. Thus there will be no extinction in finite time. Thus as long as the new plant survives past the initial year of life and both the carrying capacity and below critical harvesting is maintained, the threat of extinction is avoided.

We further noted that an increase in carrying capacity linearly increases the critical harvest. So for high yield, one could think of enhancing water and soil fertility. This provokes further study on the effects of this inputs on carrying capacity hence the yield.

41

4.6 Harvesting a logistic population proportional to size

In the previous section we discussed constant harvest on a logistic population. In this case we suppose that the number harvested is a constant fraction of the population size. Here we consider that the harvest function is $H_t = qEN_t$ where $q \geq 0$ is the harvestability coefficient defined as a fraction of the population harvested by a unit of effort E, $E \geq 0$ is the harvesting effort which is a measure of human activity to extract members of the population. In our case we consider q as the probability of correctly detecting or finding a mature culm. Equation (4.22) implies that the harvest per unit effort is a function of the population size.

$$\frac{H_t}{E} = qN_t \tag{4.22}$$

The logistic differential equation of proportional to size harvest is given by

$$\frac{dN_t}{dt} = rN_t \left[1 - \frac{N_t}{K}\right] - qN_t E \tag{4.23}$$

The rest of notations take previous meanings. Here we note that the larger the population the more the effort to find mature culms.

Let the effort of harvesting be given by

$$E(t, N_t) = \alpha - \frac{\beta}{N_t}\frac{dN_t}{dt} \tag{4.24}$$

where $\alpha, \beta > 0$ are constants, be the effort function then the rate of change of the population is

$$\frac{dN_t}{dt} = rN_t \left[1 - \frac{N_t}{K}\right] - \lambda qN_t \left(\alpha - \frac{\beta}{N_t}\frac{dN_t}{dt}\right) \tag{4.25}$$

Rearranging equation (4.25) we have

$$\frac{dN_t}{dt} = \frac{r}{c}N_t \left(1 - \frac{N_t}{K}\right) - \frac{\lambda q\alpha N_t}{c} \tag{4.26}$$

where $c = (1 - \lambda q\beta)$.

We have introduced a parameter λ to the model to take care of demand for culms. Notice that when there is no demand for culms $\lambda = 0$, $c = 1$ and equation (4.26) is the usual logistic model. Similarly if we can't detect a mature culm (this happens between year 0 and 2 from planting) $q = 0$ and $c = 1$ thus no harvest is achieved.

Next we study the equilibrium behavior of system (4.26). That is at $\frac{dN_t}{dt} = 0$.

Therefore

$$\frac{-rN_t^2}{cK} + \frac{r}{c}\left[1 - \frac{\lambda q\alpha}{r}\right]N_t = 0$$
$$N_t^2 - K\left[1 - \frac{\lambda q\alpha}{r}\right]N_t = 0 \tag{4.27}$$

From (4.27) we find that $E_1 = K[1 - \lambda q\alpha/r]$ and $E_2 = 0$ are the roots. $N_t[E_2 = 0]$ is unstable while $N_t[E_1]$ is stable if $E_1 > 0$ that is when $K[1 - \lambda q\alpha/r] > 0$ or $r > \lambda q\alpha$. When $r = \lambda q\alpha$, $E_1 = 0$ as was case E_2.

Next we find the solution of equation (4.26). Let $K\left[1 - \frac{\lambda q\alpha}{r}\right] = D$ then $\frac{dN_t}{dt} = \frac{rN_t}{Kc}[D - N_t]$.

Separating the variables we have

$$\frac{dN_t}{(N_t[D - N_t])} = \frac{r}{Kc}dt \tag{4.28}$$

The LHS of equation (4.28) is decomposable to partial fractions so that

$$\frac{1}{D}\left[\frac{1}{N_t} + \frac{1}{D - N_t}\right]dN_t = \frac{r}{Kc}dt$$

Integrating both sides we have $\ln N_t - \ln|D - N_t| = \frac{Drt}{Kc} + A$ for some constant A determined by including initial conditions $N_t[t = 0] = N_0$. We get $A = \ln\left|\frac{N_0}{D - N_0}\right|$. So

$$\ln\left|\frac{N_t}{D - N_t}\right| = \ln\left|\frac{N_0}{D - N_0}\right| + \frac{Drt}{Kc} \tag{4.29}$$

Taking exponents of both sides of equation (4.29) we have

$$\frac{N_t}{D - N_t} = \left[\frac{N_0}{D - N_0}\right]e^{\frac{Drt}{Kc}}$$

From which

$$N_t = \frac{\left[\frac{N_0}{D - N_0}\right]De^{\frac{Drt}{Kc}}}{1 + \left[\frac{N_0}{D - N_0}\right]e^{\frac{Drt}{Kc}}}$$

This is simplified to

$$N_t = \frac{N_0 D}{N_0 + [D - N_0]e^{\frac{-rDt}{Kc}}} \tag{4.30}$$

Reverting to initial notation where $D = K\left(1 - \frac{\lambda q\alpha}{r}\right)$ and $c = (1 - \lambda q\beta)$

The explicit solution is

$$N_t = \frac{N_0 K \left(1 - \frac{\lambda q \alpha}{r}\right)}{N_0 + \left(K \left(1 - \frac{\lambda q \alpha}{r}\right) - N_0\right) e^{\frac{-r\left(1 - \frac{\lambda q \alpha}{r}\right)t}{(1 - \lambda q \beta)}}} \tag{4.31}$$

where λ is the fraction of harvestable culms on demand, q is the probability of finding a mature culm in the clump α and β are constants that determine effort required to harvest. r is the intrinsic population growth rate. $N_0 = N_t(t = 0)$ and K is the environmental carrying capacity.

Model equation (4.31) was simulated for 30 years with the growth rate r determined from actual data as follows:

Let $\lambda = 0$ or $q = 0$ so that equation (4.31) is the logistic growth without harvest [see equation (4.19)] From data, $N_0 = 1, K = 25, \overline{N_1} = 4$ culms approximately. Replacing these in equation (4.20), and making r the subject we have

$$r = -\frac{1}{t} \ln \left| \frac{25}{24 N_t} - \frac{1}{24} \right| \tag{4.32}$$

From this we get $r = 1.56868$ as the intrinsic growth rate.

It is desirable that no harvesting is done within the first two years since there is no mature culm in the clump. This is also advised by Ongugo [64]. So we need to include in this model a delay like before so that there is pure growth within that period and harvesting commences at year 3. Our model that takes care of this scenario reads (equation (4.33)). This is equivalent to holding $q = 0$ until $t \geq 3$ years.

$$\frac{dN_t}{dt} = \begin{cases} rN_t \left(1 - \frac{N_t}{K}\right), & 0 \leq t < d \\ \frac{r}{c} N_t \left(1 - \frac{N_t}{K}\right) - \frac{\lambda q \alpha}{c} N_t, & t \geq d \end{cases} \tag{4.33}$$

$d = 3$ the harvest delay to maturity of culms. Figure 4.4 shows the simulated paths taken by logistic growth without harvest and with harvest at various levels of q, the proportion of mature culms in a clump. In all cases, we have assumed that demand for bamboo is equal to all available mature culms. That is $\lambda = 1$. The equation for number of culms at

any time t is

$$
N_t = \begin{cases}
\dfrac{KN_0}{N_0 + (K - N_0)e^{-rt}}, & 0 \leq t < d \\[4ex]
\dfrac{N_0 K \left(1 - \frac{\lambda q \alpha}{r}\right)}{N_0 + \left(K \left(1 - \frac{\lambda q \alpha}{r}\right) - N_0\right) e^{\frac{-r\left(1 - \frac{\lambda q \alpha}{r}\right)t}{(1 - \lambda q \beta)}}}, & t \geq d
\end{cases}
\tag{4.34}
$$

Here the critical harvest is at $1 > \frac{\lambda q \alpha}{r}$ which implies $r > \lambda q \alpha$. Since q is the probability of finding a mature culm, it can be assumed that $q < 1$ almost surely. By Markov chain analysis [9] gave 60% as the proportion of culms that are mature at age $d = 3$ years. Further, assuming that all mature culms will be on demand λ, letting $\alpha = \beta = 1$, we maintain the population at equilibrium when $\lambda q = \frac{r}{2}$ at which

$$
N_{eq} = \frac{N_0 K \left(1 - \frac{1}{2}\right)}{N_0 + \left(K \left(1 - \frac{1}{2}\right) - N_0\right) e^{\frac{-r\left(1 - \frac{1}{2}\right)t}{(1 - \frac{r}{2})}}} = \frac{\frac{N_0 K}{2}}{\left(N_0 + \left(\frac{K}{2} - N_0\right) e^{\frac{-rt}{(2-r)}}\right)}
\tag{4.35}
$$

Further $Y_e q = \lambda q N_{eq} = \frac{rK}{4}$ which means that the harvest should not deplete the population below $N_{eq} = 16.34$ culms. This is equivalent to harvesting $Y(eq) = 16.34(0.6) = 9.8$ culms per clump.
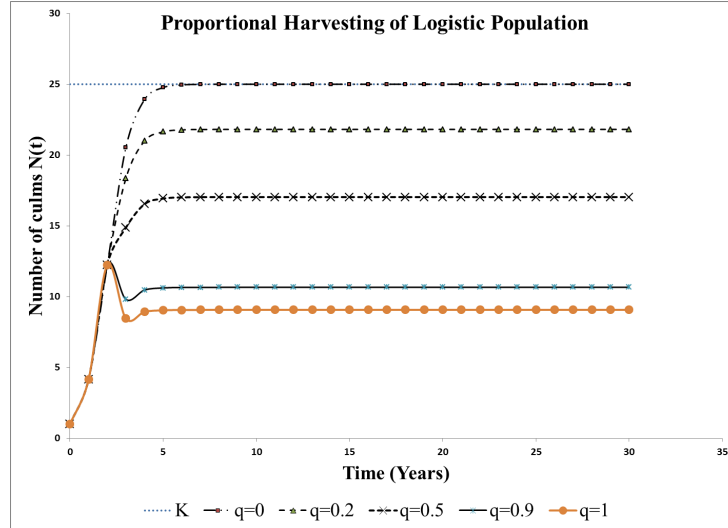
Figure 4.6: Yield proportional to population size Logistic model

The model shows that if harvest is taken proportional to the population size, some equilibrium population size lower than the carrying capacity is maintained. Simulation results shows that α can be varied within $0 < \alpha < \frac{r}{\lambda q}$ while $0 \leq \beta \leq 1$ is the domain for β. The higher the α the lower the equilibrium point since the harvest is increased.

Let N_{eq} be the equilibrium population. Table 4.1 shows the values of N_{eq} at various

instances of α with $q = 0.6$ and $\beta = \lambda = 1$.

Table 4.1: Equilibrium population at various levels of α

α	N_{eq}
1.050	20.9825
2.000	17.3500
6.535	0.7437
7.000	0.0000

α is the part of effort that does not depend on the population. So one does not care how much to harvest, the whole population may be harvested leading to extinction in finite time.

Next we investigate β and its domain. When $\beta = 0$ we have $c = 1$. Let $\lambda = 1, q = 0.6$ then $N_{eq} = 15.437$ culms.

Next suppose we increase β to $\beta = 5$; $N_{eq} = 44.5$ culms which is much greater than the carrying capacity ($K = 25$). Much closer if $\beta = 2$ we have $N_{eq} = 26.915$ which is slightly higher than the carrying capacity. We therefore restrict β to the boundaries $0 \le \beta \le 1$.

Thus our model explicitly reads

$$\frac{dN_t}{dt} = \begin{cases} rN_t \left(1 - \dfrac{N_t}{K}\right), & 0 \le t < d \\ \dfrac{r}{(1 - \lambda q \beta)} N_t \left(1 - \dfrac{N_t}{K}\right) - \dfrac{\lambda q \alpha}{(1 - \lambda q \beta)} N_t, & t \ge d \end{cases} \quad (4.36)$$

with $0 < \alpha < \frac{r}{\lambda q}, 0 \le \beta \le 1, N_0 = 1, 0 < \lambda < 1, 0 < q < 1$

4.7 Stochastic logistic modeling

We have already characterized the deterministic logistic model equation (4.1) whose solution N_t is given in equation (4.2). We saw that it has stationary solutions at $N_{eq} = 0$ and $N_{eq} = K$.

Since population dynamics in nature are usually affected by environmental variations and human activity, factors that cannot be measured and accurately modelled, stochastic equations are inevitable. Our interest is to derive a stochastic logistic model with the two harvesting situations: non - volatile and stochastic harvesting.

4.7.1   Stochastic growth example

Let the birth rate r be stochastic such that $r_t = \bar{r}_t + \sigma W_t$ where $\bar{r}_t = \mu$ a constant birth rate. Consider a stochastic logistic growth equation given by;

$$dN_t = (\mu N_t - N_t^2)dt + \sigma N_t dW_t \tag{4.37}$$

W_t is the standard Brownian process with zero mean and variance t. The associated Fokker-Planck equation is:

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial N_t}\left[\left(N_t^2 - \mu N_t + \frac{\sigma^2}{2}\frac{\partial}{\partial N_t}N_t^2\right)P\right] \tag{4.38}$$

If there is a stationary state then it must satisfy

$$0 = \frac{\partial}{\partial N_t}\left[\left(N_t^2 - \mu N_t + \frac{\sigma^2}{2}\frac{\partial N_t^2}{\partial N_t}\right)P^{st}\right] \tag{4.39}$$

This is generally written as

$$\frac{\partial}{\partial N_t}\left[\left(-f + \frac{1}{2}\frac{\partial}{\partial N_t}g(x)^2\right)P^{st}\right] \tag{4.40}$$

for which we find;

$$P^{st}(N_t) = \mathcal{L}N_t^{2\left(\frac{\mu}{\sigma^2}-1\right)}e^{-\frac{2}{\sigma^2}N_t} \tag{4.41}$$

We use the stochastic property of total probability equals unit to find the normalization constant $\mathcal{L}$.

$$1 = \int P^{st}(N_t)dN_t = \mathcal{L}\left(\frac{\sigma^2}{2}\right)^{\left(\frac{2\mu}{\sigma^2}-1\right)}\Gamma\left(\frac{2\mu}{\sigma^2}-1\right) \tag{4.42}$$

Where $\Gamma(\alpha)$ is the gamma function,hence

$$\mathcal{L} = \frac{\left(\frac{\sigma^2}{2}\right)^{-\left(\frac{2\mu}{\sigma^2}-1\right)}}{\Gamma\left(\frac{2\mu}{\sigma^2}-1\right)} \tag{4.43}$$

We note that for $\frac{\mu}{\sigma^2} \leq \frac{1}{2}$. $\mathcal{L}$ is not bounded so there is no stationary state for $N_t > 0$.

4.7.2   Stochastic logistic growth with stochastic harvesting

There are many forms of the stochastic logistic models depending on which part of the system is assumed stochastic. Many researchers prefer using the system described by equation (4.44) where only the linear part of the equation is associated with noise see

for example [4, 65, 88, 85, 53, 99, 54, 83, 49, 90] and many other texts we reviewed but did not cite in our work. No explanation as to why, or a discussion of how superior the model is to our suggested model (4.45) is given.

$$dN_t = rN_t \left(1 - \frac{N_t}{K}\right) + \sigma N_t dW_t, \quad t \geq 0 \tag{4.44}$$

We claim that both the linear and density term in N_t significantly contribute to the growth rate so that the logistic stochastic model is of the form

$$dN_t = N_t \left(1 - \frac{N_t}{K}\right)(rdt + \sigma dW_t), \quad t \geq 0 \tag{4.45}$$

Here the growth rate is stochastic. The drift can be separated from the Weiner process so that

$$dN_t = N_t \left(1 - \frac{N_t}{K}\right) rdt + \sigma N_t \left(1 - \frac{N_t}{K}\right) dW_t \tag{4.46}$$

Just like in the deterministic case equation (4.1), equation (4.45) has two equilibrium points, 0 and K. Moreover, it has been shown Jingliang and Wang [49] that equation (4.44), that is, the solution of the SDE satisfies:

$$0 < \liminf_{t \to \infty} (N_t)_i < \limsup_{t \to \infty} (N_t)_i < \infty \text{ almost surely for } i = 1, 2, 3, \ldots$$

We are equally confident that (4.45) is stochastically permanent. Alvarez and Shepp [4] showed that for Equation (4.44) N_t neither reaches zero nor infinity in finite time, and provided $> r\frac{\sigma^2}{2}$ the process has been shown to have a stationary distribution. They further showed that the probability density of N_t is

$$\chi_n^2 \left(\frac{4r}{K\sigma^2}\right) N_t \tag{4.47}$$

with continuous density

$$P(N_t) = \frac{\left(\frac{4r}{K\sigma^2}\right)^{\frac{n}{2}} 2^{-\frac{n}{2}} N_t^{\frac{n-2}{2}} e^{-\frac{4rN_t}{2K\sigma^2}}}{\Gamma\left(\frac{n}{2}\right)} \tag{4.48}$$

where $\frac{n}{2} = \left(\frac{2r}{\sigma^2} - 1\right)$

They also gave the solution for equation (4.45) as

$$N_t = \frac{N_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma B_t\right)}{1 + \left(\frac{N_0}{K}\right) r \int_0^t \exp\left(\left(r - \frac{1}{2}\sigma^2\right) s + \sigma B_s\right) ds} \tag{4.49}$$

In our work, we determine a stochastic logistic harvesting equation in which both the growth and harvest term are stochastic, They evolve by equation (4.45). We proceed to solve the equation and discuss the characteristics. Since the population growth rate is given by r, then harvest rate λ should be subtracted from the growth rate so that the net growth rate, says $\rho = r - \lambda$. We now drop assumptions (i) and (iv) in section (4.1) so that we have the stochastic term $\rho_t = r_t - \lambda_t$. Essentially, each parameter has a drift that is constant and a varying part with simple Weiner processes as

$$r_t = r + \sigma_1 \frac{dW}{dt} \tag{4.50}$$

$$\lambda_t = \lambda + \sigma_2 \frac{d\mathcal{W}}{dt} \tag{4.51}$$

Here W and $\mathcal{W}$ are independent Weiner processes representing noise in growth and harvesting respectively, σ_1 and σ_2 are noise intensities of the two parameters respectively. The net change in population is therefore

$$\rho_t = (r - \lambda) + \sigma_1 \frac{dW}{dt} - \sigma_2 \frac{d\mathcal{W}}{dt} \tag{4.52}$$

$$\rho_t = \rho + \sigma_1 \frac{dW}{dt} - \sigma_2 \frac{d\mathcal{W}}{dt} \tag{4.53}$$

Replacing r with ρ_t in equation (4.46) we have

$$dN_t = N_t \left(1 - \frac{N_t}{K}\right)\rho dt + \sigma_1 N_t \left(1 - \frac{N_t}{K}\right) dW + \sigma_2 N_t \left(1 - \frac{N_t}{K}\right) d\mathcal{W} \tag{4.54}$$

We transform equation (4.54) using

$$\ln \left| \frac{K - N_t}{N_t} \right| = f(N_t, t)$$

Which implies

$$\frac{\partial f}{\partial N_t} = \frac{-K}{N_t(K - N_t)}, \frac{\partial^2 f}{\partial N_t^2} = \frac{K(K - 2N_t)}{N_t^2(K - N_t)^2}, \text{ and} \frac{\partial f}{\partial t} = 0$$

, by Itô's Lemma, we have

$$d\ln \left| \frac{K - N_t}{N_t} \right| = \frac{\partial f}{\partial t}[dt] + \frac{\partial f}{\partial N_t}[dN_t] + \frac{\partial^2 f}{\partial N_t^2}[dN_t]^2 \tag{4.55}$$

After replacing dN_t with equation (4.54) and simplifying we have

$$d\ln \left| \frac{K - N_t}{N_t} \right| = -\rho dt + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)dt - \frac{N_t}{K}(\sigma_1^2 + \sigma_2^2)dt - \sigma_1 dW + \sigma_2 d\mathcal{W} \tag{4.56}$$

Note that we have used same procedure as the example equation (3.14) we solved as an application of Itô's Lemma (Section 3.2.8). For simplicity of expression, let $(\sigma_1^2 + \sigma_2^2) = \mathcal{V}$ then integrating on both sides in the range $[0, t)$ we have

$$\ln\left|\frac{K - N_t}{N_t}\right| - \ln\left|\frac{K - N_0}{N_0}\right| = \int_0^t \left[\left(\left(\rho - \frac{1}{2}\mathcal{V}\right) + \frac{N_s}{K}\mathcal{V}\right) ds + \sigma_1 dW - \sigma_2 d\mathcal{W}\right]$$

Taking exponents of both sides and simplifying we get

$$N_t = \frac{K N_0}{N_0 + (K - N_0)e^{A(t)}} \tag{4.57}$$

Where

$$A_t = -\int_0^t \left[\left(\left(\rho - \frac{1}{2}\mathcal{V}\right) + \frac{N_s}{K}\mathcal{V}\right) ds + \sigma_1 dW - \sigma_2 d\mathcal{W}\right)$$

We note that $\left(\rho - \frac{1}{2}\mathcal{V}\right)$ is constant with respect to time so that

$$A_t = -\left(\rho - \frac{1}{2}\mathcal{V}\right) t - \int_0^t \left(\left[\frac{N_s}{K}\mathcal{V}\right] ds + \sigma_1 dW - \sigma_2 d\mathcal{W}\right) \tag{4.58}$$

Here $\frac{N_s}{K}$ at any instance is a fraction multiplied by $\mathcal{V}$ which is total variance, and coupled with noisy variations in growth and harvest rates. For sufficiently little noise, $dW \to 0$ and $d\mathcal{W} \to 0$ so that

$$A_t = -\left(\rho - \frac{1}{2}\mathcal{V}\right) t - \frac{N_t}{K}\mathcal{V}t = -\rho t + \left(\frac{1}{2} + \frac{N_t}{K}\right)\mathcal{V}t \tag{4.59}$$

We can now examine the extreme regimes. Clearly $\lim_{t \to \infty} A_t \to -\infty$ and equation (4.57) maximizes at $N_t = K$. So however large the noise, the population does not explode to infinity. This basically agrees with nature that however how much fertilizer, rain and other growth accelerators we apply; there are a maximum number of individuals that an ecosystem can carry. At the lower end. When $t = 0$ we have $A_t = 0$ and $N_t = N_0$. Further, we analyze harvest and possibilities of extinction. If A_t is negative, the population grows larger. While a positive A_t diminishes the population.

The only time A_t will be positive is when $-\rho + \left(\frac{1}{2} + \frac{N_t}{K}\right)\mathcal{V} > 0$ or when $-(r - \lambda) + \left(\frac{1}{2} + \frac{N_t}{K}\right)\mathcal{V} > 0$ or $\lambda > r - \left(\frac{1}{2} + \frac{N_t}{K}\right)\mathcal{V}$. This means that regardless the amount of "good noise" (for example, a good amount of unexpected rainfall), large harvest can drive a population to extinction. Let K be much larger compared to N_t such that $\frac{N_t}{K} = 0$, then the condition for persistence of the population is that

$$\lambda < r - \frac{1}{2}\mathcal{V} \text{ or } \lambda < r - \frac{1}{2}(\sigma_1^2 + \sigma_2^2)$$

Comparing this with the deterministic equivalent, the difference is in the variance. We also note that even without harvesting, a sufficiently large noise can drive the population to extinction. From equations (4.57) and (4.58), derivation of harvest equations when either harvest or growth rate are not stochastic is straight forward. The difference is only in the variance where we let either of the two (σ_1^2 or σ_2^2)equal to zero appropriately. The form of the solution is the same with change only in A_t. A simulation of the system at various levels of noise is plotted in the following figures.

Simulated populations under noise intensities $\sigma_1 = 0.1, \sigma_2 = 0.1$ with various levels of average proportional harvesting rate h are given in the following set of graphs (Figure 4.7). In our model we were able to relax various assumptions so that the system evolves as naturally as possible. Constant growth and constant harvest rates are no-longer mandatory. We also imposed the delay to onset of harvesting, a factor that can be easily noted as the graphs break abruptly at 3 years.

The graphs (in Figure 4.7) bring out the seriousness of how many culms are cut within a time period. The higher the harvest, the worse the equilibrium status of the system. As the harvest rate approaches growth rate, the system approaches extinction in finite time.
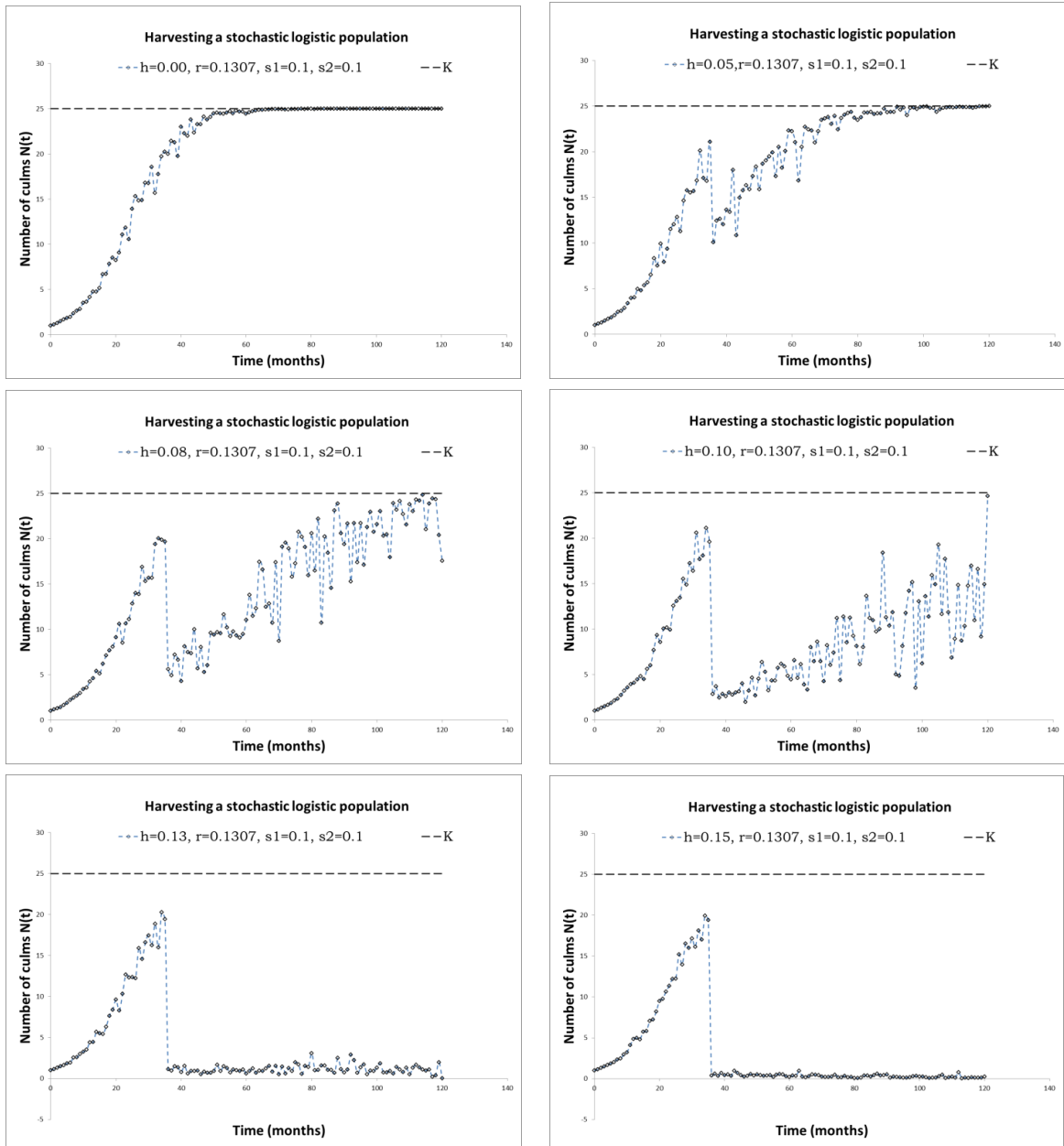
Figure 4.7: Harvesting a stochastic logistic population in a real Brownian environment

Finally, maximum optimal yield was simulated at $\sigma_1 = \sigma_2 = 0.1$ noise levels. Figure 4.8 shows a maximum average sustainable yield of 15.5 culms per clump. That is at $h = 0.07$, indeed this is almost half of the growth rate $r = 0.1307$. The variation is due to the system noise.
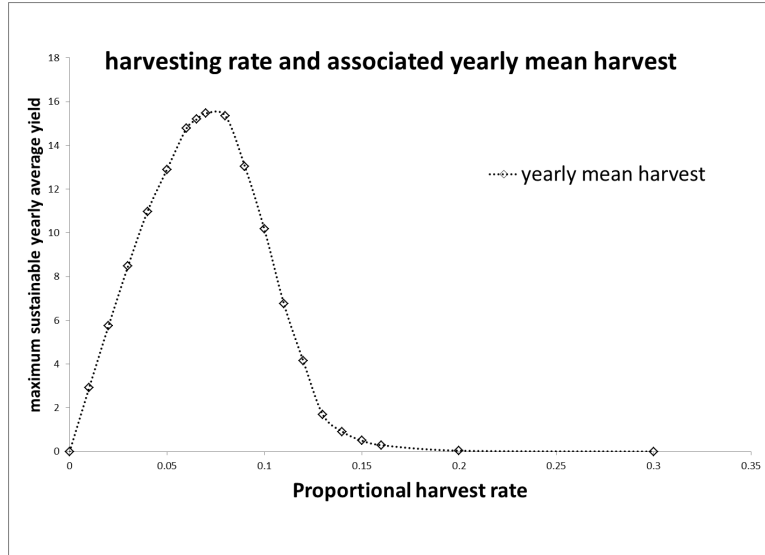
Figure 4.8: Harvesting rate and associated yearly mean harvest

The optimum harvest strategy for a fully stochastic system has been derived. We have also shown that the maximum sustainable yield and some level of noise is averagely half the growth rate. The critical harvesting rate is therefore given as

$$\lambda = \frac{1}{2}\left(r - \frac{1}{2}(\sigma_1^2 + \sigma_2^2)\right) \qquad (4.60)$$

A graph of optimal strategy together with associated mean yearly yield (dark dotted line) is given (Figure 4.9). The light dotted line is a no harvest trajectory. The graph brings out a conclusion that with sufficient noise that does not cause system extinction, the system struggle to head to carrying capacity. Since harvest is proportional to population size, and if there is infinite demand for the culms, then with confidence we can say that there is no absolute steady state in a noisy system other than extinction. We are only sure that the system is bounded between Zero and K. This has been noted in the deterministic cases as well.
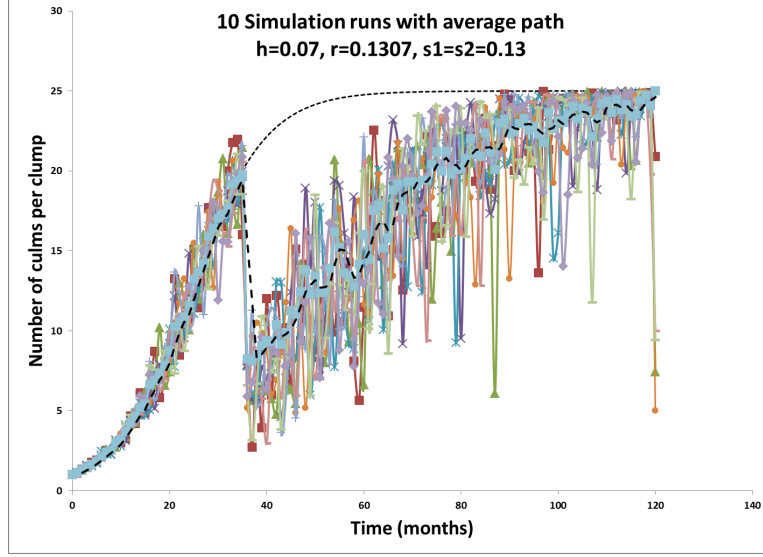
Figure 4.9: 10 simulation runs of the stochastic logistic harvest model with noise in growth and harvesting

4.7.3 Stochastic logistic model with linear term random

We consider the model of the form

$$dN_t = [(r - \lambda)N_t - \beta N_t^2]dt + \sigma_1 N_t dW_t^{(1)} - \sigma_2 N_t dW_t^{(2)} \tag{4.61}$$

in which λ is the harvesting rate per population unit. Here the net growth is only multiplied by the linear term in N_t. We had earlier under the deterministic harvest modeling (section 4.6) given the proportional harvesting rate as $qEN_t = H_t$. Thus here $\lambda = qE$. We have relaxed the assumptions of constant growth rate and harvest so that both are stochastic. σ_1 and σ_2 are the noise intensities for growth rate and harvest rate respectively. Equation (4.61) can be written as

$$dN_t = m(N_t)dt + \sigma(N_t)dW_t \tag{4.62}$$

where

$$m(N_t) = [(r - \lambda)N_t - \beta N_t^2], \quad \sigma(N_t) = (\sigma_1 N_t, -\sigma_2 N_t) \quad \text{and} \quad W_t = \left(W_t^{(1)}, W_t^{(2)}\right)^T$$

Equation (4.62) is a continuous Markov process for which the probability density function of the solution can be determined using the corresponding Fokker-Planck equation.

$$\frac{\partial P(N_t, t, \lambda)}{\partial t} = \frac{-\partial}{\partial N_t}M(N_t)P(N_t, t, \lambda) + \frac{\partial^2}{2\partial N_t^2}[\sigma^2(N_t)P(N_t, t, \lambda)] \tag{4.63}$$

Since the drift $M(N_t)$ and the noise term $\sigma(N_t)$ are invariant with respect to time, the equation that satisfies the stationary distribution is

$$\frac{d}{d(N_t)}[M(N_t)P(N_t, \lambda)] - \frac{\partial^2}{2\partial N_t^2}(\sigma^2(N_t)P(N_t, \lambda)) = 0 \tag{4.64}$$

From the generalization given in the stochastic growth example (Section 4.7.1), (4.64) has an explicit solution in the domain $[0, +\infty)$. That is

$$P(N_t, \lambda) = \mathcal{L}N_t^{2\left(\frac{r-\lambda}{\sigma_1^2+\sigma_2^2}\right)-2} e^{-2\left(\frac{\beta N_t}{\sigma_1^2+\sigma_2^2}\right)} \tag{4.65}$$

where the normalization constant $\mathcal{L}$ is found using the integral

$$\int_0^\infty \mathcal{L}N_t^{2\left(\frac{r-\lambda}{\sigma_1^2+\sigma_2^2}\right)-2} e^{-2\left(\frac{\beta N_t}{\sigma_1^2+\sigma_2^2}\right)} dN_t = 1 \tag{4.66}$$

When $\lambda < \left(r - \frac{\sigma_1^2+\sigma_2^2}{2}\right)$ is true we have

$$\mathcal{L} = \left(\int_0^\infty N_t^{\left(\frac{2r-2\lambda}{\sigma_1^2+\sigma_2^2}\right)-2} e^{-\left(\frac{2\beta N_t}{\sigma_1^2+\sigma_2^2}\right)} dN_t\right)^{-1} \tag{4.67}$$

So

$$P(N_t, \lambda) = \frac{N_t^{\left(\frac{2r-2\lambda}{\sigma_1^2+\sigma_2^2}\right)-2} e^{-\left(\frac{2\beta N_t}{\sigma_1^2+\sigma_2^2}\right)}}{\int_0^\infty N_t^{\left(\frac{2r-2\lambda}{\sigma_1^2+\sigma_2^2}\right)-2} e^{-\left(\frac{2\beta N_t}{\sigma_1^2+\sigma_2^2}\right)} dN_t} \tag{4.68}$$

For simplicity of expression let $\alpha = \frac{2r}{\sigma_1^2+\sigma_2^2}$ and $\theta = \frac{2}{\sigma_1^2+\sigma_2^2}$ then (4.68) becomes

$$P(N_t, \lambda) = \frac{N_t^{(\alpha-\theta\lambda)-2} e^{-\theta\beta N_t}}{\int_0^\infty N_t^{(\alpha-\theta\lambda)-2} e^{-\theta\beta N_t)dN_t}} \tag{4.69}$$

The expectation of sustainable yield is computed as

$$\overline{Y}_s = \mathbb{E}(N_t, \lambda) = \int_0^\infty \lambda N_t P(N_t, \lambda) dN_t$$

$$\overline{Y}_s = \frac{\lambda \int_0^\infty N_t^{(\alpha-\theta\lambda)-1} e^{-\theta\beta N_t} dN_t}{\int_0^\infty N_t^{(\alpha-\theta\lambda)-2} e^{-\theta\beta N_t)dN_t}} \tag{4.70}$$

Let $\theta\beta N_t = y$ then $dy = \theta\beta dN_t$ and $N_t = \frac{y}{\theta\beta}$ we have

$$\overline{Y}_s = \frac{\lambda \int_0^\infty y^{(\alpha-\theta\lambda)-1}e^{-y}dy}{(\theta\beta)\int_0^\infty y^{(\alpha-\theta\lambda)-2}e^{-y}dy} = \frac{\lambda}{\theta\beta}\frac{\Gamma(\alpha-\theta\lambda)}{\Gamma(\alpha-\theta\lambda-1)}$$
$$= \frac{\lambda}{\theta\beta}\frac{\lambda(\alpha-\theta\lambda-1)\Gamma(\alpha-\theta\lambda-1)}{\Gamma(\alpha-\theta\lambda-1)} = \frac{\lambda(\alpha-\theta\lambda-1)}{\theta\beta} \tag{4.71}$$

Replacing α and θ we have

$$\overline{Y}_s = \frac{\lambda\left(\frac{2r}{\sigma_1^2+\sigma_2^2}\right)-\left(\frac{2\lambda}{\sigma_1^2+\sigma_2^2}\right)-1}{\frac{2\beta}{\sigma_1^2+\sigma_2^2}} = \frac{\lambda[2(r-\lambda)-(\sigma_1^2+\sigma_2^2)]}{2\beta} \tag{4.72}$$

Equation (4.72) can be reorganized so that

$$\overline{Y}_s = \frac{\lambda[(r-\lambda)-\frac{1}{2}(\sigma_1^2+\sigma_2^2)]}{\beta} \tag{4.73}$$

From this we note that if $(r-\lambda) = \frac{\sigma_1^2+\sigma_2^2}{2}$ we have no yield. Thus, $(r-\lambda) > \frac{\sigma_1^2+\sigma_2^2}{2}$ or $\lambda < \left(r-\frac{\sigma_1^2+\sigma_2^2}{2}\right)$ is a necessary condition for any harvest to be realized. Considering conditions previously prescribed at equilibrium, (Section 4.5, condition 1 and section 4.6 with the condition that follows equation (4.35) in the deterministic case, we have the optimal harvesting policy at $\lambda = \frac{1}{2}\left(r-\frac{\sigma_1^2+\sigma_2^2}{2}\right)$.

$$\max \lambda_s = \left(\frac{r}{2} - \frac{\sigma_1^2+\sigma_2^2}{4}\right) \tag{4.74}$$

Substituting (4.74) in (4.73) we have a maximum sustainable yield of

$$\max\overline{Y}_s = \frac{1}{4\beta}\left[r - \frac{\sigma_1^2+\sigma_2^2}{2}\right]^2 \tag{4.75}$$

Its variance can be shown to be

$$Var\overline{Y}_s = \frac{\sigma_1^2+\sigma_2^2}{16\beta^2}\left[r - \frac{\sigma_1^2+\sigma_2^2}{2}\right]^3 \tag{4.76}$$

where r is the intrinsic rate and $\beta = \frac{r}{K}$ with K the carrying capacity. We note that the population neither reaches some steady state nor goes extinct but cycles between some maximum and minimum boundaries determined by the level of noise in the two rates.

4.7.4 Stochastic logistic model with only growth rate being random

We have considered two cases (sections 4.7.2 and 4.7.3) where both the growth rate and harvest are stochastic. This is the most realistic scenario than other variants of the stochastic logistic family of models. It is easily verifiable that when the randomness assumption is not taken into account (by letting $\sigma_1 = \sigma_2 = 0$) our results coincide with those we derived under deterministic logistic model of the Verhulst population. If the farmer is organized and a strictly proportional yield is taken, then $\sigma_2 = 0$ and the following results hold.

i. Equation (4.62) will read

$$dN_t = [(r - \lambda)N_t - \beta N_t^2]dt + \sigma_1 N_t dW_t^{(1)} \tag{4.77}$$

ii. The necessary condition to avoid extinction reads

$$\max \lambda_s = \left(\frac{r}{2} - \frac{\sigma_1^2}{4} \right) \tag{4.78}$$

iii. The probability distribution of the system is

$$P(N_t, \lambda) = \frac{N_t^{2\left(\frac{r-\lambda}{\sigma_1^2}\right)-2} e^{-2\left(\frac{\beta N_t}{\sigma_1^2}\right)}}{\int_0^\infty N_t^{2\left(\frac{r-\lambda}{\sigma_1^2}\right)-2} e^{-2\left(\frac{\beta N_t}{\sigma_1^2}\right)} dN_t} \tag{4.79}$$

iv. The maximum yield

$$\max \overline{Y}_s = \frac{1}{4\beta} \left[r - \frac{\sigma_1^2}{2} \right]^2 \tag{4.80}$$

v. Variance of yield

$$Var \overline{Y}_s = \frac{\sigma_1^2}{16\beta^2} \left[r - \frac{\sigma_1^2}{2} \right]^3, \text{ for } r > \frac{\sigma_1^2}{2} \tag{4.81}$$

The maximum sustainable yield at various levels of noise in growth rate together with the associated variance of yield are plotted in Figure 4.10. The higher the noise, the lower the sustainable harvest.
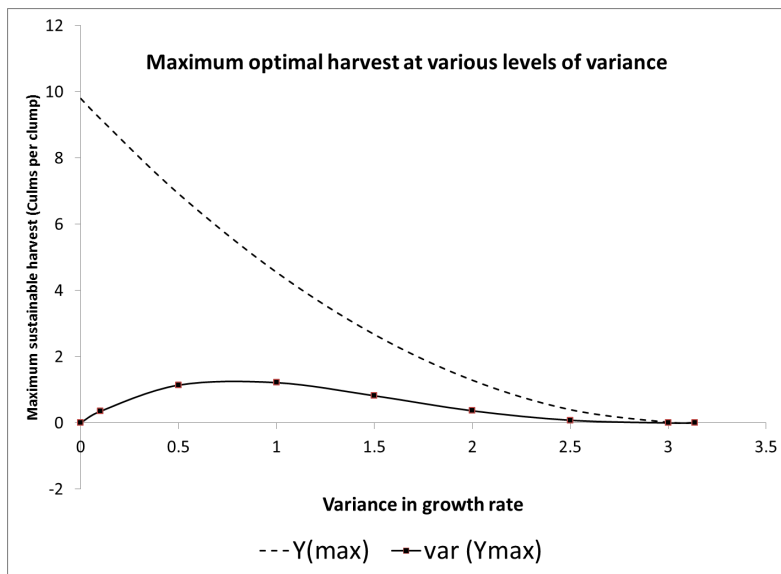
Figure 4.10: Optimum harvest at various levels of harvest

CHAPTER FIVE

RESULTS AND DISCUSSION

5.1 Introduction

In this chapter, computations and findings on the following aspects of bamboo growth are reported. Growth in culm height, culm diameter, clump circumference, number of culms per clump and crown cover. These are all important in this research since they give a complete description of the plants under investigation. Further, multiple regression models are fitted to data and their goodness of fit discussed. The bamboo growth variables are modeled by linear regression analysis and prediction equations of one in relation to another determined. Sites are analyzed to determine which species of bamboo does best in each location. Climatic data for the study sites is analyzed and used to predict yield. Hypotheses formulated in the study are also tested. A 5-parameter logistic model is introduced and used in determining yield. The model simulated output is compared with actual data for model verification. Further, the deterministic and stochastic models discussed in chapter 4 are applied and their impact in the prediction of yield, together with the advantage of equilibrium harvesting are discussed.

5.2 Growth in culm height

Growth in height was monitored by measuring the tallest three culms per monitored clump. Note that not the same culms in tagged clumps were measured over time. Each new shoot is normally thicker and grows taller than previous ones, only taking around 3 months [50] to stretch to maximum height. After this, maturity takes over two years.

5.2.1 *Bambusa vulgaris* height analyses

Average heights for the two species in each site, taken at various points in time were as summarized in Table 5.1.

Table 5.1: Mean growth in height for bambusa vulgaris in the four study sites

Period (Months)	Bambusa μ (Suna E.)	Bambusa μ (Rangwe)	Bambusa μ (Suba)	Bambusa μ (Kuria)
3	0.4845	0.4428	0.4229	0.5127
6	1.4535	1.3284	1.2687	1.5380
9	2.1944	2.3396	1.6837	2.2514
12	3.1047	3.3407	2.4797	2.9525
19	4.2722	4.5515	3.7276	4.6487
26	5.2242	5.9416	4.4963	5.7321
36	6.3208	7.2637	5.8500	6.7850

Here μ is the mean height (in meters) attained by bamboo at the given site. For easier comparison, the data were fitted to polynomial regression curves of order 3. These were found to be best fitting to the data as shown in Figure 5.1.
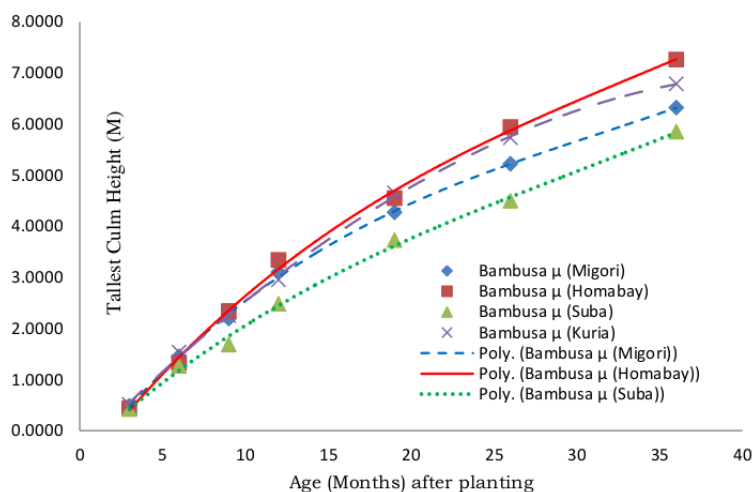
Figure 5.1: Growth in culm height for bambusa vulgaris at the four study sites

Equations for growth in height (m) for b*ambusa vulgaris* species.

Rangwe:

$$y = 0.000082x^3 - 0.008309x^2 + 0.416580x - 0.778544$$
$$R^2 = 0.998048 \qquad x \geq 3 \text{ months}$$

Suna E.

$$y = 0.000101x^3 - 0.009506x^2 + 0.405122x - 0.661444$$

$$R^2 = 0.999262 \qquad x \geq 3 \text{ months}$$

Kuria:

$$y = -0.000009x^3 - 0.003190x^2 + 0.325679x - 0.396447$$

$$R^2 = 0.999243 \qquad x \geq 3 \text{ months}$$

Suba:

$$y = 0.000065x^3 - 0.005877x^2 + 0.301060x - 0.429473$$

$$R^2 = 0.997529 \qquad x \geq 3 \text{ months}$$

In all cases the goodness of fit was plausible since, as read from values of the coefficient of determination R^2, age x in months explains over 99% of variation in height y (meters).

5.2.2 Giant bamboo height analyses

The height of a bamboo culm contributes proportionately to the harvest volume. *dendrocalamus giganteus* bamboos are known to get to a maximum height of 35M while *bambusa vulgaris* reaches up to a maximum of 20M [72]. Table 5.2 shows the mean height of giant bamboo at various time intervals corresponding to monitoring times.

Table 5.2: Mean growth in height for *dendrocalamus giganteus* in the four study sites

Period (Months)	Giant μ (Suna E.)	Giant μ (Rangwe)	Giant μ (Suba)	Giant μ (Kuria)
3	0.4871	0.4905	0.5195	0.4355
6	1.4612	1.4714	1.5586	1.3065
9	2.5591	2.8202	2.5228	2.3768
12	3.3145	3.7810	3.1969	3.2675
19	4.6456	4.9791	3.9001	5.4511
26	5.6586	6.8417	5.0720	6.5024
36	7.6350	9.9379	7.4083	8.9333

The data were fitted with polynomial regression curves for easy comparison (Figure 5.2). The ranking of sites in terms of average giant bamboo height performance was Rangwe, Kuria, Suna E., then Suba in descending order.
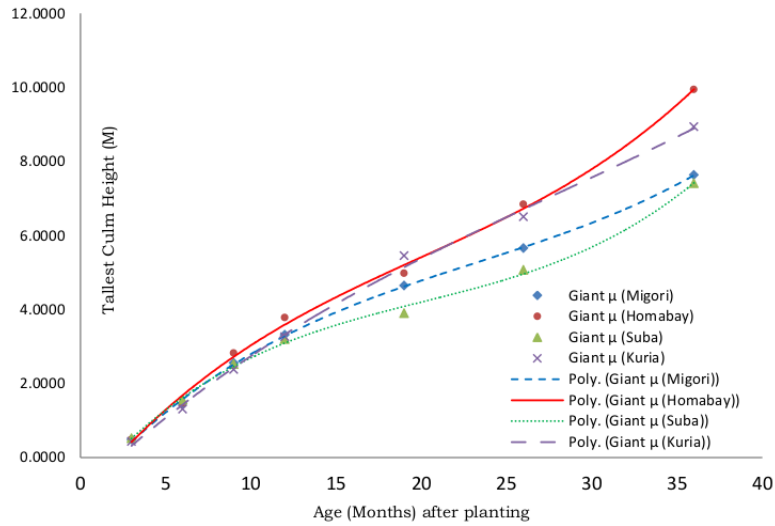
Figure 5.2: Growth in culm height for *dendrocalamus giganteus* at the four study sites

In each site, growth was sigmoid. The best fit model was polynomial order 3 which had the best coefficient of determination ($R^2 > 0.99$). The following are site specific equations for growth in height for *dendrocalamus giganteus* species. In all cases the condition $x \geq 3$ holds.

Rangwe:

$$y = 0.000294x^3 - 0.017617x^2 + 0.561131x - 1.109421$$
$$R^2 = 0.997515$$

Kuria:

$$y = 0.000094x^3 - 0.007933x^2 + 0.435899x - 0.917744$$
$$R^2 = 0.996893$$

Suna E.:

$$y = 0.000223x^3 - 0.015552x^2 + 0.508966x - 0.962877$$
$$R^2 = 0.999452$$

Suba:

$$y = 0.000344x^3 - 0.020744x^2 + 0.532380x - 0.899151$$
$$R^2 = 0.997960$$

Again in all cases the goodness of fit was plausible since time(x) explains over 99% of variation in height. The general trend for growth in height for the two species in the study region is summarized in Table 5.3 with a graphical representation in Figure 5.17.

Table 5.3: Mean general growth in height for *dendrocalamus giganteus* and *bambusa vulgaris*

Period (Months)	Bambusa height (M)	Giant height (M)
3	0.4657	0.4831
6	1.3971	1.4494
9	2.1173	2.5697
12	2.9694	3.3900
19	4.3000	4.6940
26	5.3486	5.8436
36	6.5549	8.6786

Again, due to the genetic composition, *dendrocalamus giganteus* had higher heights each monitoring time, over the study period.
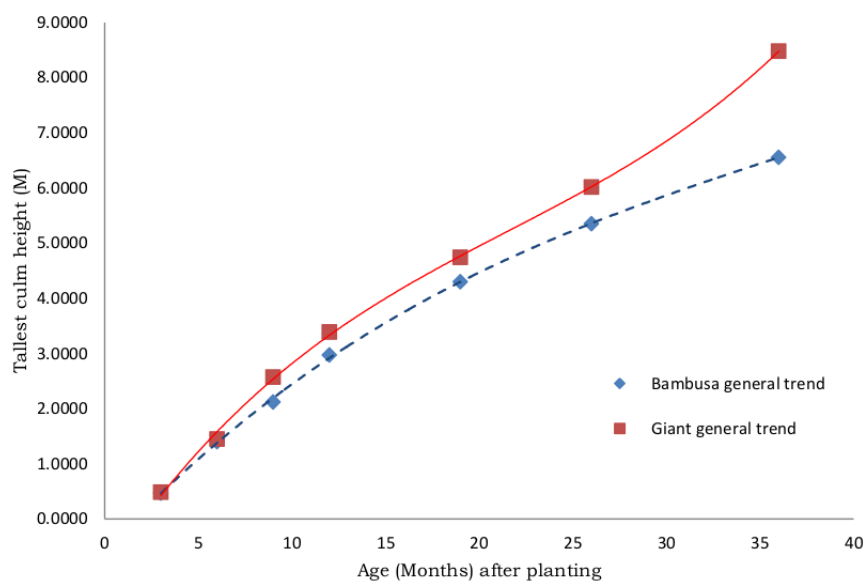
Figure 5.3: Mean general growth in height for *d. giganteus* and *b. vulgaris*

The equations for general trend are:

Giant bamboo *(d. giganteus)*

$$y = 0.000239x^3 - 0.015461x^2 + 0.509594x - 0.972298$$
$$R^2 = 0.999433, \quad x \geq 3$$

bambusa vulgaris

$$y = 0.000060x^3 - 0.006720x^2 + 0.362111x - 0.566477$$
$$R^2 = 0.999687, \quad x \geq 3$$

The trend shows a slow rise in height with time, with *dendrocalamus giganteus* exuding a slightly faster growth rate than *bambusa vulgaris* with time. This behavior was replicated in all the sites.

The general average height at age 36 months was 6.555 M for *bambusa vulgaris* and 8.679 M for *dendrocalamus giganteus*. The maximum single height was 17.867 M for *bambusa vulgaris* found in Rangwe farm H17 code B12 owned by Joseph Nondi. The tallest *dendrocalamus giganteus* culm was 17.867 M tall. This was also found in Rangwe H12 owned by Samson Juma.

5.3   Clump diameter

This is the diameter of a whole cluster of poles emerging from one seedling. The clump diameter was computed by measuring the circumference of a tagged cluster and multiplying by $1/\pi$. Measures of clump circumference were taken at age 26 and 32 months as shown in the following Figure.
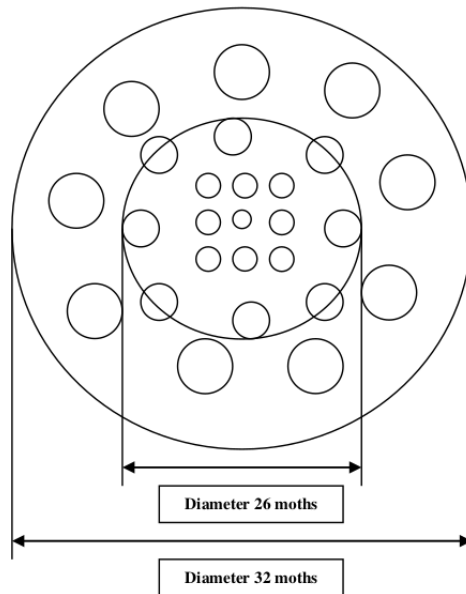
Figure 5.4: Growth in culm and clump diameter with time.

Data on clump diameter from the four sites was analyzed to compare growth among the

two species. In comparing giant and *bambusa* clump diameters, the question was if there is significant difference in clump diameter of the two species.

Table 5.4: Mean clump diameter for the *d. giganteaus* and *b. vulgaris* species

Period (months)	Site	CLUMP DIAMETER Giant (M)	CLUMP DIAMETER *Bambusa* (M)
26	Rangwe	0.502	0.658
26	Suna E.	0.496	0.855
26	Kuria	0.513	1.012
26	Suba	0.459	0.615
32	Rangwe	0.790	0.966
32	Suna E.	0.758	1.109
32	Kuria	0.850	1.348
32	Suba	0.686	0.848

For all 167 Giant and 244 *bambusa vulgaris* consistently investigated clumps for each of the two species, the following clump diameters data were obtained.

Table 5.5: Mean clump diameter from a sample of clumps

	Giant clumps	*bambusa* clumps
Sample size	167	244
Mean clump diameter (M)	0.63175	0.986125
Standard deviation	0.156189217	0.262832069

The correlation coefficient of average clump diameter over the four sites between giant and *bambusa vulgaris* clumps was 0.926, meaning that there was a strong relationship between clump diameters among the two species at the different sites. Further a test of hypothesis about the difference between means of the two species gives Z (computed) as -5.18, and the critical Z (tabulated) at 5% level of significance (two sided) is 1.96

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

where μ is mean clump diameter for *d. giganteus* and μ for *b. vulgaris.* The p-value 0.0254 was <0.05. Further, since the absolute value of the test statistic Z was greater than 1.96, we rejected the null hypothesis and concluded that there was a significant difference in the mean clump diameters of the two bamboo species. The negative sign on the test statistic value shows that Giant bamboo had lesser clump diameters than the bambusa species at the given points in time. Indeed, bambusa culms spread out faster than giant culms [72].

5.4 Crown diameter

Crown diameter was measured by measuring the widest and narrowest spread of each clump and averaging. The study of crown cover is important since as the cover increases, the possibility of intercropping diminishes. Light penetration to the ground is also hindered hence inhibiting growth of any other crop within the clumps. This also affects further shooting of bamboo hence lowering the yield.

Table 5.6: Average crown diameter for *d. giganteaus* and *b. vulgaris*

	BAMBUSA CROWN DIAMETER (M)			GIANT CROWN DIAMETER (METERS)		
	WIDEST	NARROWEST	AVERAGE	WIDEST	NARROWEST	AVERAGE
SUBA	3.643	1.200	2.327	3.453	1.299	2.471
KURIA	4.223	1.402	3.238	5.075	1.470	2.847
SUNA E.	4.006	1.358	2.669	4.180	1.398	2.702
RANGWE	3.949	1.308	2.713	4.119	1.341	2.645

For the two species planted at the same time, the relationship between crown diameters was investigated. A correlation analysis of the averages in Table 5.6 gave 0.962 implying a strong positive relationship.

A test of the null hypothesis that there is no significant difference between mean *bambusa* crown diameters and mean giant crown diameter against the alternative of inequality was made:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

Where μ_1 is mean crown diameter for *d. giganteus* and μ_2 for *b. vulgaris*, gave $\overline{X_1} = 2.737$M (*d. giganteus*) and $\overline{X_2} = 2.667$M (*b. vulgaris*) with $n_1 = 167$ (number of giant clumps investigated) and $n_2 = 244$ (number of *bambusa* clumps investigated). Standard deviation for the giant clump diameter $S_1 = 0.703959$ and that of *bambusa* is $S_2 = 0.597487$.

The computed Z value was 1.05355 and the critical value at 5% level of significance 1.96. This led to failure to reject the null hypothesis leading to the conclusion that there was no significant difference in crown diameter between the two species.

5.5 Culm diameter

One of the most significant measures in bamboo is culm diameter. Technically, it is measured at breast height hence the name diameter at breast height (dbh). Dbh data were recorded from 3 largest culms in each tagged clump. Note that culm diameter for a particular mature culm never changes with time. It was earlier illustrated (Figure 5.4) that new culms have larger diameters and are further from the clump center than previous ones. The mean culm diameters for the two species over the study period are presented in the following tables.

Table 5.7: Mean culm diameter (cm) for *b. vulgaris* at the four sites

period (Months)	Rangwe	Suna E.	Kuria	Suba
3	0.4	0.245376	0.453754	0.195376
6	0.7992553	0.710336711	0.84735117	0.6564809
9	1.3981381	1.407778	1.437747	1.348138
12	2.092399	1.818687	1.840225	1.768687
19	2.400525	2.757754	2.3	2.25
26	3.7466667	3.877987	4.241667	3.696667
36	4.6706349	4.746667	4.871667	4.620635

The mean culm diameter at age 36 months in all sites was between 4.6 and 4.9 cm. No signficant difference was noted in this dimension to suppose one sight as superior to another.

Table 5.8: Mean culm diameter (cm) for *d. giganteus* at the four sites

period (Months)	Rangwe	Suna E.	Kuria	Suba
3	0.7	0.56375	0.6	0.51375
6	1.1269069	0.8969	0.85619	0.78444
9	1.7672673	1.396667	1.240476	1.190476
12	2.2338624	1.930354	1.983175	1.880354
19	3.0110184	3.675654	3.3	2.961018
26	5.4855556	5.178616	5.885714	5.128616
36	7.8924242	6.825	7.765	6.775

For *d. giganteus*, mean culm diameter was different from site to site. It ranged between 6.7cm and 7.9cm thick. Data in tables 5.8 and 5.9 were fitted multiple with regression curves for visual comparison. The plots are as shown in Figures 5.5 and 5.6, respectively.
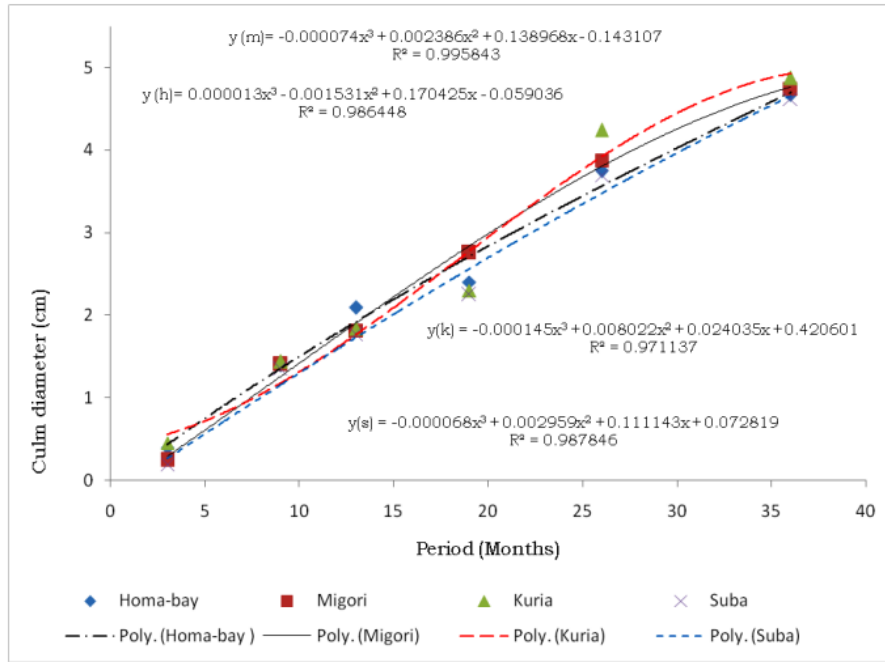
Figure 5.5: A polynomial fit for mean culm diameter for *b. vulgaris* at the four sites
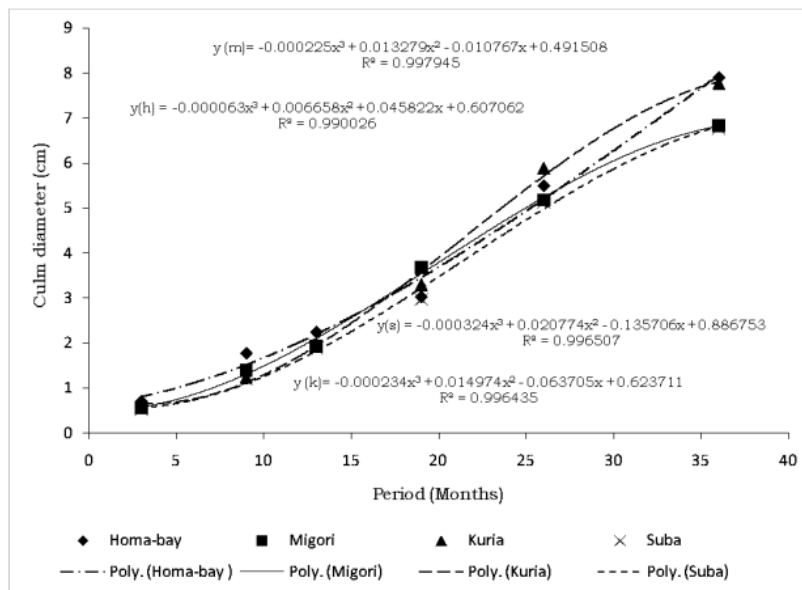
Figure 5.6: A polynomial fit for mean culm diameter for *d. giganteus* at the four sites

In both graphs, all regression equations have the condition $x \geq 3$. As was the case with culm height, growth in culm diameter was also sigmoid. A polynomial of order three gave a good fit with the coefficient of determination $R^2 > 0.99$. Average culm diameters over time for each species were tabulated as follows.

Table 5.9: Mean culm diameter for the two species over time at the four study sites

Period (Months)	*D. giganteus* mean culm diameter	*B. vulgaris* mean culm diameter
3	0.62125	0.366377
6	0.916114	0.75335599
9	1.468137	1.414554
12	2.04913	1.917104
19	3.328891	2.486093
26	5.516629	3.95544
36	7.494141	4.682156

To visualize the trend for the data, a polynomial fit for the data were plotted (Figure 5.7). At the start, 3 to 12 months, both species had similar diametric growth rates. Thereafter, *dendrocalamus giganteus* had significantly larger diameters than *bambusa vulgaris* at all sites.
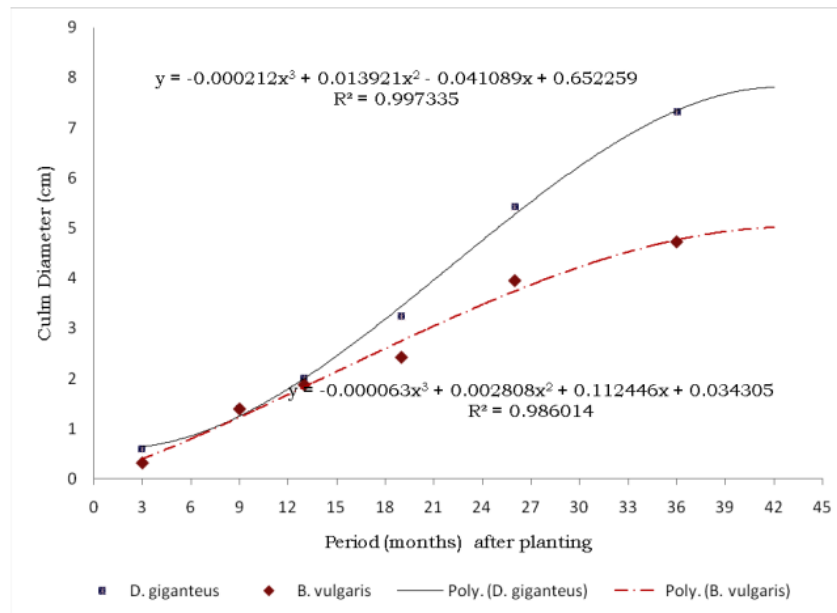
Figure 5.7: Mean general growth in culm diameter for *b. vulgaris* and *d. giganteus*

5.6 Comparing height and culm diameter

The heights and corresponding culm diameters of the best three culms per tagged clump were recorded. Data for mean measures of the two parameters were as in Table 5.10. There was a strong relationship between height and culm diameter as seen from Table 5.11.

Table 5.10: Height and culm diameter for giant bamboo

Table 10: Height and culm diameter for giant bamboo

Period (Months)	Giant height (M)	Giant culm diameter (cm)
3	0.48314325	0.594375
6	1.4494295	0.916113631
9	2.569733	1.398721579
12	3.38999225	2.006936027
19	4.7439765	3.236922765
26	6.01864575	5.419625637
36	8.47863625	7.314356061

Table 5.11: Correlations between height and culm diameter

Correlations for height vs culm diameter	Suna E. (giant)	Homa-bay (giant)	Suba (giant)	Kuria (giant)	General trend
d- giganteus	0.9809921	0.983866021	0.97162647	0.973055	0.98038855
b-vulgaris	0.9932834	0.992090933	0.98496776	0.9718988	0.98831247

There was a strong linear relationship between giant culm diameter and height of each culm. Using the regression equation, as height tends to a maximum of 30M, the culm diameter tends to 30 cm wide. The linear regression equation for giant bamboo is:

$$y(\text{culm height M}) = 1.0743x(\text{culm diameter cm}) + 0.6707 \quad R^2 = 0.9612$$

Comparing height against culm diameter for *bambusa vulgaris* also showed a strong positive linear relationship with a steeper slope than that of *dendrocalamus giganteus*. Using the regression equation, it is estimated that as culm height tends to 20M, culm diameter tends to 15 cm as expected.
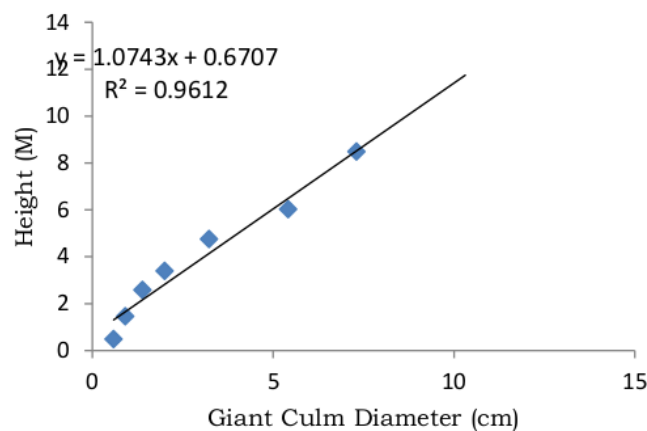
Figure 5.8: A linear graph of height against culm diameter for giant bamboo

$x \geq 3$ for the linear equation in Figure 5.8. The linear model is however not accurate on

extrapolation. For *b. vulgaris*, the following table has average height and corresponding average culm diameter.

Table 5.12: Height and culm diameter for *bambusa vulgaris*

Period (Months)	*Bambusa* height (M)	*Bambusa* culm diameter (cm)
3	0.4657	0.3236
6	1.3971	0.7534
9	2.1173	1.3980
12	2.9694	1.8800
19	4.3000	2.4271
26	5.3486	3.8907
36	6.5549	4.7274

Equation for *bambusa vulgaris*.

$$y(\text{culm height}) = 1.3439x(\text{culm diameter}) + 0.3509, \quad R^2 = 0.9768$$

We can therefore predict height given culm diameter using the linear regression model. From table 5.13 for *bambusa vulgaris*, both the intercept and gradient were significant at 5% level. The 95% confidence interval for the gradient was (1.1056, 1.5822)
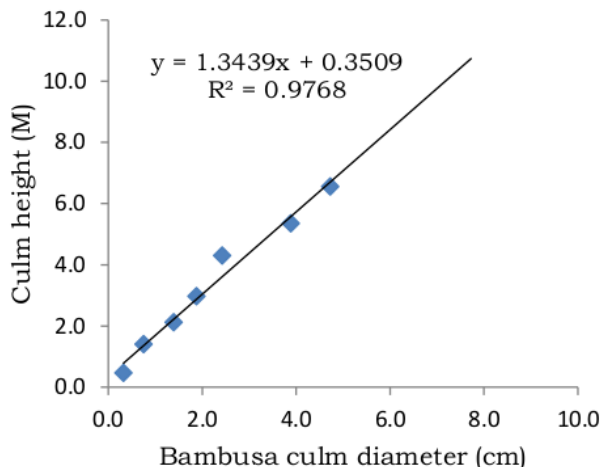
Figure 5.9: A graph of height against culm diameter for *bambusa vulgaris*

The regression statistics are given in the following table.

Table 5.13: Regression coefficients for height against culm diameter for *b. vulgaris*

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.350868	0.246617	1.42272	0.2140953	-0.28308	0.984820
culm diameter (cm)	1.343942	0.092705	14.4969	2.818E-05	1.10563	1.582249

From Table 5.14 for *dendrocalamus giganteus*, the gradient was significant at 5% level. The 95% confidence interval for the gradient was (0.8260, 1.3225).

Table 5.14: Regression coefficients for height against culm diameter for *d. giganteus*

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.670691	0.365982	1.832582	0.126347	-0.27009	1.611477
Giant culm diameter (cm)	1.074289	0.096576	11.1238	0.000102	0.826033	1.322544

bambusa vulgaris had a steeper gradient than *dendrocalamus giganteus* implying that the lateral growth in *bambusa vulgaris* is less prominent than longitudinal growth. The converse was true for *dendrocalamus giganteus*.

5.7   Growth in number of culms

The number of culms per clump is an important measure of the productivity of a farm. Depending on soil and weather conditions, there are varied amounts of culms a clump carries at a given time from the date of planting.

Data on the number of culms per clump was recorded at various time points up to the harvest time (3 years from planting). This was used to plot the following growth curves (Figure 5.10, 5.11 and 5.12).
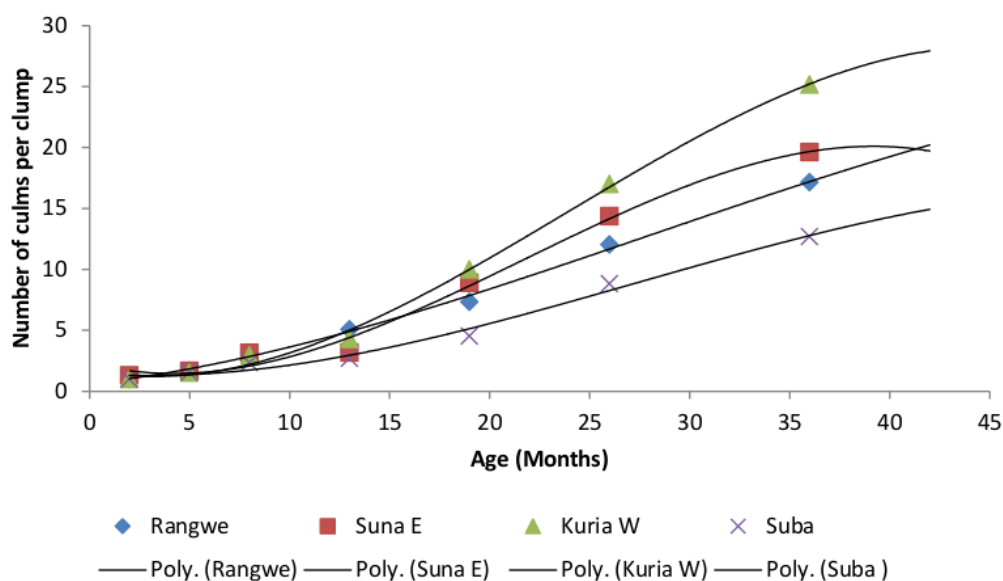
Figure 5.10: Growth in number of *bambusa vulgaris* culms per clump at the four study sites

The polynomial equations for growth in number of culms per clump for each site are as follows:

Kuria W.:

$$y = -0.0008x^3 + 0.0548x^2 - 0.3308x + 1.7409$$
$$R^2 = 0.9975, \quad x \geq 3$$

Suna E.:

$$y = -0.0009x^3 + 0.0566x^2 - 0.4277x + 2.3202$$
$$R^2 = 0.9903, \quad x \geq 3$$

Rangwe:

$$y = -0.0002x^3 + 0.0136x^2 + 0.1818x + 0.614$$
$$R^2 = 0.9974, \quad x \geq 3$$

Suba:

$$y = -0.0003x^3 + 0.0217x^2 - 0.1236x + 1.4798$$
$$R^2 = 0.989, \quad x \geq 3$$

From (Figure 5.10 there was clear difference in site performance in terms of the average number of *bambusa* culms per clump observed with time. The same observation was made in the case of giant bamboo at the sites (Figure 5.11).
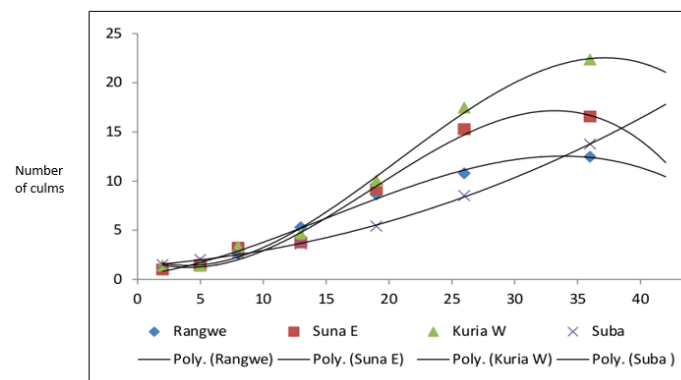
Figure 5.11: Growth in number of *dendrocalamus giganteus* culms

The polynomial equations for growth in number of culms per clump in each site for Giant bamboo are:

$$y(\text{Kuria W}) = -0.0012x^3 + 0.0714x^2 - 0.5158x + 2.4173$$
$$R^2 = 0.9952, \quad x \geq 3$$
$$y(\text{Suna E}) = -0.0013x^3 + 0.0723x^2 - 0.5268x + 2.2507$$
$$R^2 = 0.9874, \quad x \geq 3$$
$$y(\text{Rangwe}) = -0.0005x^3 + 0.0222x^2 + 0.1636x + 0.4082$$
$$R^2 = 0.9965, \quad x \geq 3$$
$$y(\text{Suba}) = 2E - 05x^3 + 0.0064x^2 + 0.0911x + 1.3679$$
$$R^2 = 0.9993, \quad x \geq 3$$

The general trend for both species was plotted in Figure 5.12. No clear difference was noted in terms of the rate of production of culms for both species in the early period of development. However later, the multiplicity of culms of *b. vulgaris* became more rampant than that of *dendrocalamus giganteus.*
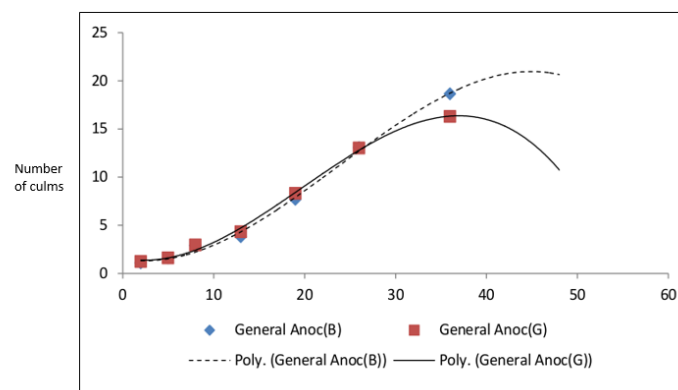
Figure 5.12: Growth in number of culms per clump for the two species

The general polynomial equations for the two species are:

$$y\,(bambusa) = -0.0005x^3 + 0.0367x^2 - 0.1751x + 1.5387$$
$$R^2 = 0.9963, \quad x \geq 3$$
$$y\,(Giant) = -0.0007x^3 + 0.0431x^2 - 0.197x + 1.611$$
$$R^2 = 0.9974, \quad x \geq 3$$

The poles in each tagged clump were counted at the end of three years from the date of planting. This was done for all poles including mature and immature ones. Descriptive statistics for the number of culms per clump for *bambusa vulgaris* at the four sites are as shown in table 5.15.

Table 5.15: General statistics for *b. vulgaris* number of culms per clump

B. vulgaris number of culms	Rangwe	Kuria	Suna E.	Suba
Mean	16.619	23.05	18.325	10.166
S Error	0.848	1.803	0.835	1.759
S Deviation	5.495	8.062	5.279	4.309
Range	24	24	21	10
Minimum	8	10	9	5
Maximum	32	34	30	15
C I (95.0%)	1.712	3.773	1.688	4.522

In this and subsequent tables, confidence interval figures are mean ± the corresponding figures in the last row. The best site for propagating *bambusa vulgaris* was Kuria with an average of 23±4 culms per clump at the time of harvesting, followed by Suna E. with 18±2 culms, then Rangwe with 17±2 culms. Suba was the worst performing for the species with only 10±5 culms. Similarly, descriptive statistics for the number of culms per clump for *dendrocalamus giganteus* at the four sites is as shown in Table 5.16.

The best site for propagating *dendrocalamus giganteus* was Kuria with an average of 22±5 culms per clump at the time of harvesting, followed by Suna E. with 17±2 culms, then Rangwe and Suba with 12.36±1.23 culms and 11.67±1.94 culms respectively. The maximum number of culms per clump was 34 for *bambusa vulgaris* while that of *dendrocalamus giganteus* was 31. Both maximums were observed in Kuria.

Table 5.16: General statistics for *d. giganteus* number of culms per clump

D. giganteus number of culms	Rangwe	Kuria	Suna E.	Suba
Mean	12.3636	22.3500	16.5500	11.6667
S Error	0.6075	2.5673	0.9240	0.8819
S Deviation	4.0295	11.4812	5.8439	3.0551
Range	18	46	23	11
Minimum	5	5	7	7
Maximum	23	31	30	18
C I (95.0%)	1.2251	5.3734	1.8690	1.9411

5.8   Clump circumference

Three years after planting, clump circumference was also measured for each tagged clump in each of the four study sites. Descriptive statistics for clump circumference in *bambusa vulgaris* was as in Table 5.17.

Table 5.17: General statistics for *b. vulgaris* clump circumference

B. vulgaris clump circumference	Rangwe	Kuria	Suna E.	Suba
Mean	4.151190476	5.125	4.64375	3.35
Standard Error	0.148036786	0.219434081	0.149541902	0.420911709
Standard Deviation	0.959388022	0.981339042	0.945786032	1.031018914
Range	4.3	3.5	3.65	2.8
Minimum	2.1	3.1	2.55	2
Maximum	6.4	6.6	6.2	4.8
Confidence Level (95.0%)	0.298966351	0.459280808	0.302477044	1.081987993

Again the ranks in terms of clump circumference for the sites are the same as was the case with number of culms. The best performing was Kuria with an average of 5.13±0.46 M of clump circumference at the time of harvesting. Suba was the worst performing for the species with only 3.35±1.08M circumference. The large variability in Suba was due to extreme performances of the growth measures. Some farms were very productive while others were close to empty. Similarly, descriptive statistics for the number of culms per clump for *dendrocalamus giganteus* at the four sites was as shown in Table 5.18.

Table 5.18: General statistics for *d. giganteus* clump circumference

D. giganteus clump circumference	Rangwe	Kuria	Suna E.	Suba
Mean	3.034090909	3.8275	3.1575	2.766666667
Standard Error	0.136396666	0.321294155	0.098637351	0.200171644
Median	3.1	3.65	3.2	2.83
Standard Deviation	0.904753129	1.436871144	0.62383738	0.693414914
Range	4	6.4	3	2.15
Minimum	1.2	2.1	1.8	1.39
Maximum	5.2	8.5	4.8	3.54
Confidence Level (95.0%)	0.275070089	0.672476394	0.199512871	0.440574817

5.9 Comparing number of culms against clump circumference

A regression analysis of number of culms per clump on clump circumference was carried out with the following results.

Table 5.19: Regression statistics for number of culms vs clump circumference (*b. vulgaris*)

Regression Statistics	
Multiple R	0.873538082
R Square	0.76306878
Adjusted R Square	0.758983759
Standard Error	3.19553802
Observations	60

The simple linear regression model had dependable explanatory power $R^2 = 0.76$. This implies that 76% of the clump circumference could be determined by the number of *bambusa* culms in a clump. The regression model from the following output (figure 5.14) is:

$$\text{number of bambusa culms} = 6.406(\text{clump circumference}) - 10.33$$

with clump circumference ≥ 2 m (the minimum observed from the fields)

A linear model for the dependence of number of culms per clump on clump circumference for *d. giganteus* was also constructed giving the OLS linear regression equation as

$$\text{number of giant culms} = 6.290(\text{clump circumference}) - 3.681 \tag{5.1}$$

with clump circumference≥ 1.2 m (the minimum observed for the species)

It also had good predictive power within the range of the data with $R^2 = 0.76$ just like the common bamboo case. Suspicious data points that could be candidates for outliers were seen on the scatter plot (Figure 5.13). This could be explained by some single culm popping up a distance away from the rest in the clump, which consequently raises circumference without proportionately raising the number of culms in the clump. The two points that lie within the prediction interval but far ahead along the line are most probably two clumps that were planted on very fertile soil than the rest, hence an accelerated growth in number of culms significantly different from the rest. The points are useful since they justify extrapolation of the regression line to some further circumference. The previous outlier on the other hand is not strictly useful since it is a pretty rare occurrence that distorts the prediction power of our model.
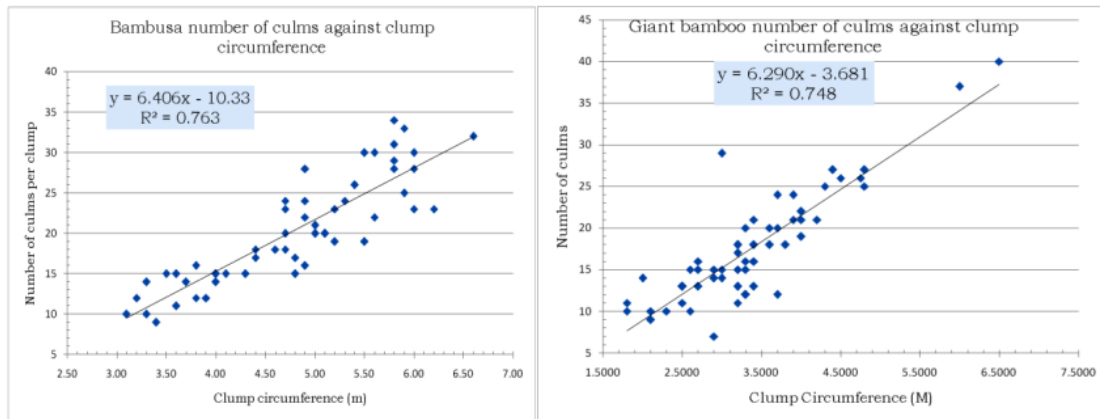
Figure 5.13: Linear regression models for number of culms against clump circumference

It is interesting to note that the gradients for the two species are approximately the same. The difference is the intercept. When the circumference is 4.5 M, for example, there are approximately 25 giant bamboo culms while there are only 17 Common bamboo culms. *b. vulgaris* has generally more sparsely populated culms than *d. giganteus*, which is a tight clumping species [50].

We should note that extrapolation of the line to estimate the number of culms for smaller circumferences is not justified and could be fallacious. When the circumference is zero, for example, the line could predict negative numbers of culms for both species.

The reliability of coefficients was tested and results (Table 5.20) show that both the gradient and intercept were significant at 5% level (both *p*-values ≪ 0.0001).

Table 5.20: Analysis of variance for goodness of fit

ANOVA

	df	SS	MS	F	Signif F
Regression	1	13.2410	13.24103	47.1354	2.012E-06
Residual	18	5.05646	0.28091		
Total	19	18.2975			

The residual plots show that residuals were generally independently identically distributed as normal with constant variance. However there was a tendancy of heteroscedasticity in the first plot. The more the number of *bambusa* culms per clump, the higher the variance. This is explained by the running tendency of *bambusa* culms as the population within a clump rises. This is prompted by search for nutrients at less populated points away from the center of the clump. There were noticeable cases of outliers in the second plot (giant bamboo). Although the plot is largely homoscedastistic (constant variance) giant bamboo takes longer to mature and there is a possibility that late in the growth process, the behavior of common bamboo would be experienced as well.

Residual plot for *b. vulgaris*

Residual plot for *d. giganteus.*
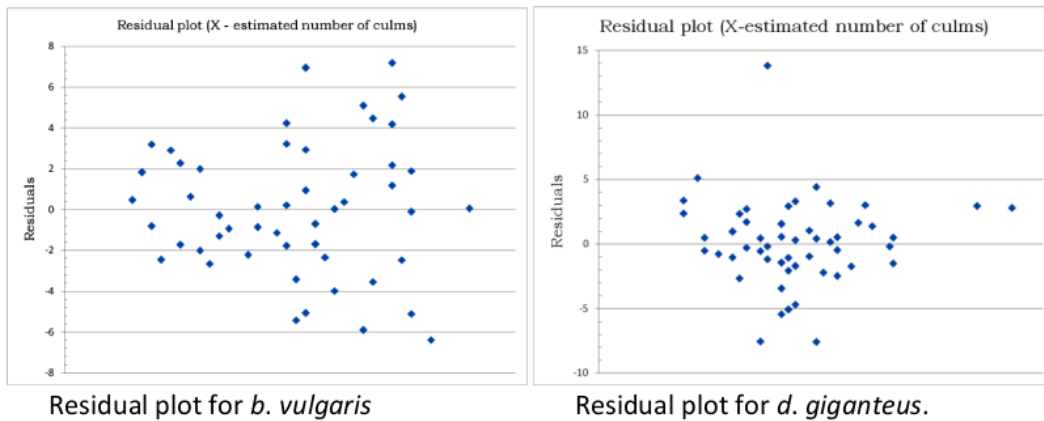
Figure 5.14: Residual plots for estimated number of culms using clump diameter

5.10 Site analysis for difference in clump circumference

The productivity of a plot of bamboo can be sufficiently measured by the clump density (number of culms per clump). Clump density depends on clump circumference and number of culms per clump. Table 5.21 shows the mean clump circumference for each site, together with associated variance.

Table 5.21: Descriptive statistics for bambusa vulgaris clump circumference per site

Groups	Average circ.(M)	Variance
Rangwe	4.15	0.920
Suna E.	4.64	0.894
Suba	3.35	1.063
Kuria	5.13	0.963

Having successfully fitted a linear regression model that predicts the number of culms given clump circumference (see Figure 5.13), it is sufficient to measure clump circumference as a predictor for harvest volume. For each species, the mean clump circumferences were compared by testing for equality of means. The hypotheses were stated as:

$$H_0 : \mu_r = \mu_m = \mu_k = \mu_s$$
$$H_1 : \text{not all } \mu_j \text{ are equal.}$$

Here subscripts correspond to names of the four sites. This is a One-Way ANOVA test that requires the following assumptions:

1. There are k simple random samples from k populations.

79

2. The k samples are independent of each other; that is, the subjects in one group cannot be related in any way to subjects in a second group.

3. The populations are normally distributed.

4. The populations have the same variance; that is, each treatment group has population variance σ^2.

Table 5.22: Analysis of variance table for equality of mean clump circumference

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	21.583	3	7.194	7.775	9.83E-05	2.692
Within Groups	96.235	104	0.925			
Total	117.819	107				

Since the p-value was much less than 0.05 H_0 was rejected, implying that there was strong evidence to show significant difference in average clump circumferences for *bambusa vulgaris* among the sites. The same verdict was arrived at when the number of culms per clump was considered as the measure of productivity.

5.11   Number of culms per clump

In checking for equality of the average number of culms per clump among the sites the following hypothesis was tested.

$$H_0 : \mu_r = \mu_m = \mu_k = \mu_s$$
$$H_1 : \text{not all } \mu_j \text{ are equal.}$$

Again, subscripts correspond to names of the four sites. We had the following means and associated variances.

Table 5.23: Descriptive statistics for *bambusa vulgaris* number of culms per clump

Groups	Average	Variance
Rangwe	16.619	30.1928
Suna E.	18.325	27.8660
Suba	10.167	18.5667
Kuria	23.050	64.9974

Table 5.24: Analysis of variance for multiple comparisons

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	961.7869	3	320.5956	9.12862	2.04E-05	2.691979
Within Groups	3652.463	104	35.11984			
Total	4614.25	107				

H_0 was again rejected with the same conclusion that there was strong evidence to show significant difference in average number of culms per clump for *bambusa vulgaris* among the sites. The one way ANOVA test justifies the conclusion that some sites were better producers of *b. vulgaris* than others. Further, the equality of means of number of *bambusa* against giant culms per clump was tested. The four null hypotheses were:

H_0: There is no significant difference in mean number of *bambusa culms* as compared to the mean number of giant culms per clump at 36 months at site S.

Table 5.25: Results from hypotheses testing

Site S	z	p-value	decision
Rangwe	-4.47828	3.76228E-06	Reject Ho
Suna E.	-2.27924	0.011326468	Reject Ho
Kuria	-0.83167	0.20279758	Fail to reject Ho
Suba	0.813692	0.792089174	Fail to reject Ho

In Kuria and Suba, there was no significant difference in the number of culms per clump. However in Rangwe and Suna E., bambusa performed significantly better than *giganteus*. Z was positive in Suba meaning *giganteus* performed better there than *bambusa*. The 95% Confidence Intervals for yield per site were as in Table 5.26.

Table 5.26: 95% confidence intervals for number of culms per clump

Species	Site	Average number of culms	N	SD	±se(95% CI)	$\overline{X_{1/36}}$ Lower limit	$\overline{X_{u/36}}$ Upper limit
Giant	Rangwe	12.4651	43	4.0199	1.201515	11.26359	13.66661
Bambusa	Rangwe	17.143	43	5.5461	1.657683	15.48532	18.80068
Giant	Suna E.	16.55	40	5.8439	1.81101	14.73899	18.36101
Bambusa	Suna E.	19.6	40	6.1218	1.897131	17.70287	21.49713
Giant	Kuria	22.35	20	11.481	5.031765	17.31823	27.38177
Bambusa	Kuria	25.15	19	9.4939	4.2689	20.8811	29.4189
Giant	Suba	13.75	20	3.304	1.448015	12.30198	15.19802
Bambusa	Suba	12.6667	25	5.5377	2.170739	10.49596	14.83744

The number of culms per clump is an important measure of the productivity of a farm. Depending on soil and weather conditions, there are varied amounts of culms a clump carries at a given time from the date of planting. Due to competition for nutrients and harvesting of mature culms, a clump cannot have infinitely many culms. The carrying

capacity of a clump varies from place to place depending on available space, soil and weather conditions. In this study, data on the number of culms per clump were recorded at various time points up to the harvest time (3 years from planting). Various models were fitted to data and their goodness of fit (GOF) checked.

5.12   Weather data and yield

5.12.1   Climate data at the study sites

Secondary climatic data for the study sites was acquired from Climate-data.org website http://en.climate-data.org/location/11136/. A summary of extracted data were tabulated and climate graphs plotted as in the following tables and figures.

Table 5.27: Climate data for various places in the study region

	Rainfall (mm)	Altitude (m)	Mean daily temp. (°C)
Kuria	1418	1634	19.6
Suna E.	1369	1366	21.2
Rangwe	1476	1314	21.6
Sindo	1033	1143	22.6
Kisii	1922	1631	19.6
Mbita	1001	1141	22.6
Homabay	1226	1166	22.5
Kisumu	1321	1178	22.9

Source (climate-data.org)

5.12.2   Suna East (Migori County) Climate

At a mean altitude of 1366 m, Migori is dominated by tropical monsoon climate. There is more rainfall in the early part of the year than in the later in Migori. The temperature here averages 21.2°C±1.9°C. Precipitation here averages 1369 mm per annum.

Migori climate 1982-2012
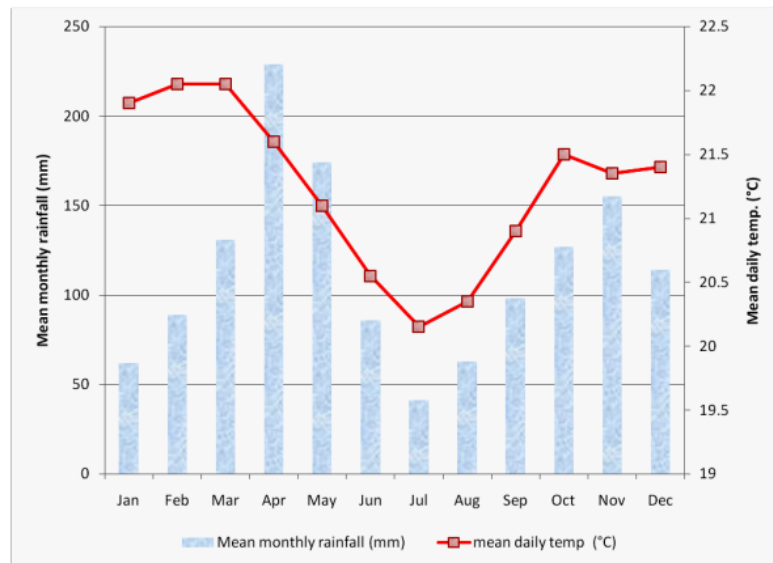
Figure 5.15: Migori climate between 1982 and 2012, Source [30]

5.12.3 Rangwe (Homabay County) Climate

At a mean altitude of 1314 m, Rangwe is also dominated by tropical climate. Rangwe has significant rainfall most months, with a short dry season. The temperature here averages 21.6°C±2.0°C. Precipitation here averages 1476 mm per annum.

Rangwe climate 1982-2012
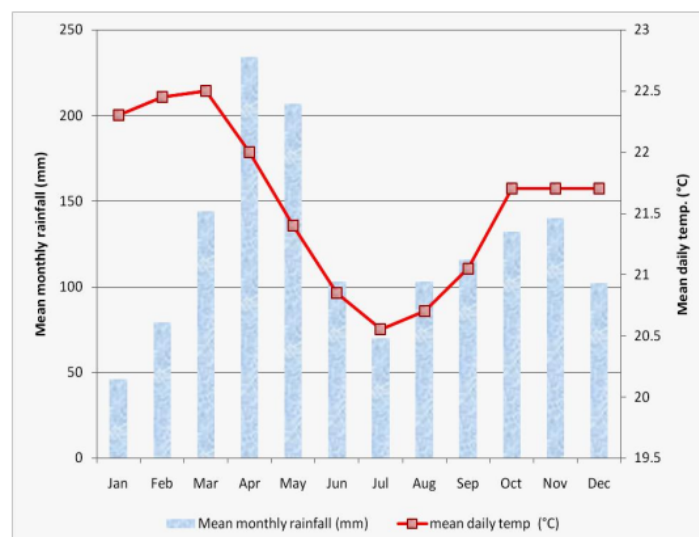
Figure 5.16: Rangwe climate between 1982 and 2012, Source [27]

5.12.4   Suba (Homabay County) Climate

Sindo in Suba Sub-county is at 1143 m altitude. Her climate is classified as tropical. The temperature here averages 22.6°C. At an average temperature of 23.5°C, February is the hottest month of the year. The lowest average temperatures in the year occur in July, when it is around 21.6°C. About 1033 mm of precipitation falls annually. Precipitation is the lowest in July, with an average of 36 mm.
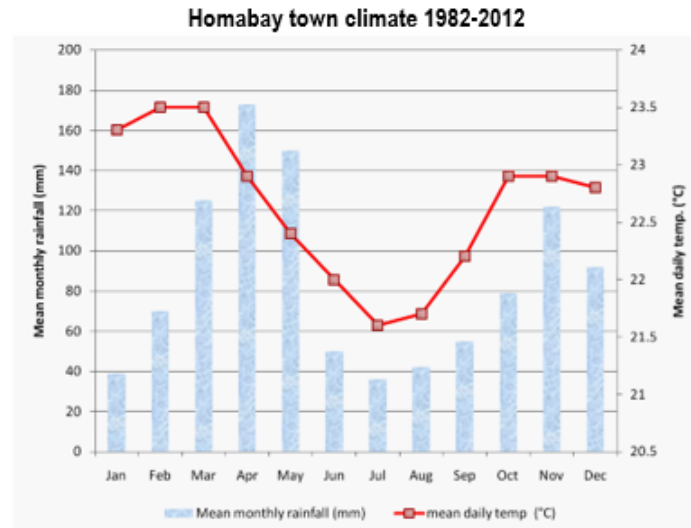
Figure 5.17: Homabay climate between 1982 and 2012, Source [28]

5.12.5   Kuria West (Migori County) Climate

Kuria is at 1634m altitude. Her climate is classified as tropical. The temperature here averages 19.6°C. The variation in annual temperature is around 1.9°C. In a year, the average yearly rainfall is at 1418 mm.

Interesting relationships were discovered between altitude, rainfall and mean temperature with high correlation coefficients. There was high positive relationship between annual rainfall and altitude of the site. Mean daily temperature had a strong negative relationship with both altitude and rainfall.
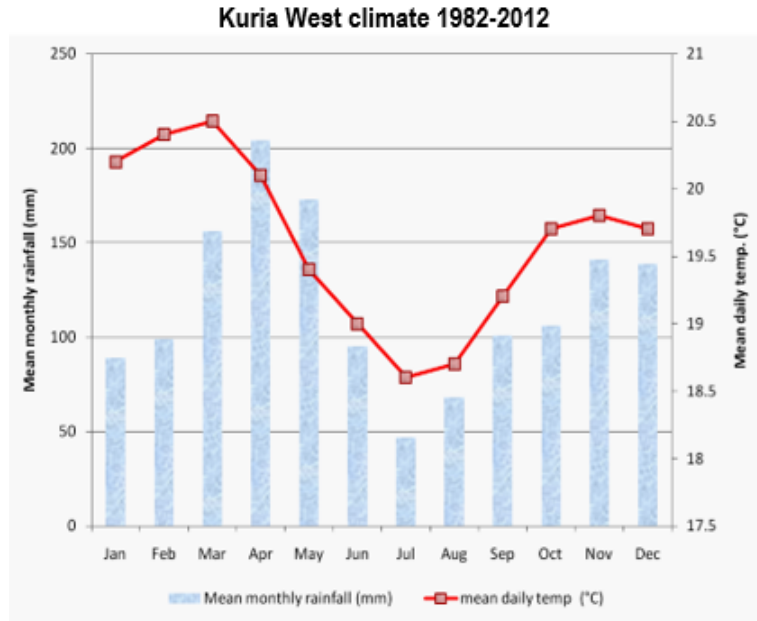
Figure 5.18: Kuria climate between 1982 and 2012, Source [29]

To validate the relationship, more regions neighboring the study sites were incorporated and the results found to be as follows.

Table 5.28: Correlation analysis for climate data

	Altitude (m)	Rainfall (mm)	Mean daily temp. (°C)
Altitude (m)	1		
Rainfall (mm)	0.807214345	1	
Mean daily temp. (°C)	-0.991657043	-0.772259565	1

This warranted a multiple regression model for forecasting rainfall as follows:

Table 5.29: Regression statistics for altitude and daily temperature on rainfall

Regression Statistics	
Multiple R	0.836374723
R Square	0.699522677
Adjusted R Square	0.579331747
Standard Error	187.724349
Observations	8

Table 5.30: ANOVA table for the multiple regression model

ANOVA

	df	SS	MS	F	Significance F
Regression	2	410205.344	205102.672	5.820095	0.049491345
Residual	5	176202.156	35240.4312		
Total	7	586407.5			

85

Table 5.31: Regression coefficients and their significance

	Coefficients	Standard Error	t Stat	P-value
Intercept	-11109.74	12298.39	-0.90	0.407753
Altitude (m)	3.46	2.64	1.31	0.247134
Mean daily temp. (°C)	365.59	409.38	0.89	0.412749

The multiple linear regression model is

$$\text{Rainfall} = -11109.74 + 3.46(\text{altitude}) + 365.59(\text{mean daily temperature})$$

The coefficient of determination (0.699) was high but the coefficients were all insignificant at 5% level. The correlation coefficient between predictors suggests existence of multi-colinearity, were the predictor variables, mean daily temperature and altitude have a high inverse correlation of -0.992. If we select the variable with stronger prediction power (Altitude) and drop the other, we have a simple linear regression situation with the following parameters.

Table 5.32: Regression statistics for altitude on rainfall

Regression Statistics	
Multiple R	0.807214345
R Square	0.651594998
Adjusted R Square	0.593527498
Standard Error	184.5297203
Observations	8

Table 5.33: ANOVA table for the simple regression model

ANOVA

	df	SS	MS	F	Significance F
Regression	1	382100.1939	382100.1939	11.22133715	0.015422689
Residual	6	204307.3061	34051.21768		
Total	7	586407.5			

Table 5.34: Regression coefficients for the simple linear regression model

	Coefficients	Standard Error	t Stat	P-value
Intercept	-134.2515	446.6054	-0.3006	0.77386
Altitude (m)	1.1198	0.3343	3.3498	0.01542

The model was

$$\text{Rainfall (mm)} = -134.25 + 1.12(\text{Altitude (m)}) \tag{5.2}$$

The coefficient of determination was 0.65 and the gradient was significant at 0.05 levels. A residual plot for the model showed no clear violation of the assumption of homoscedasticity.
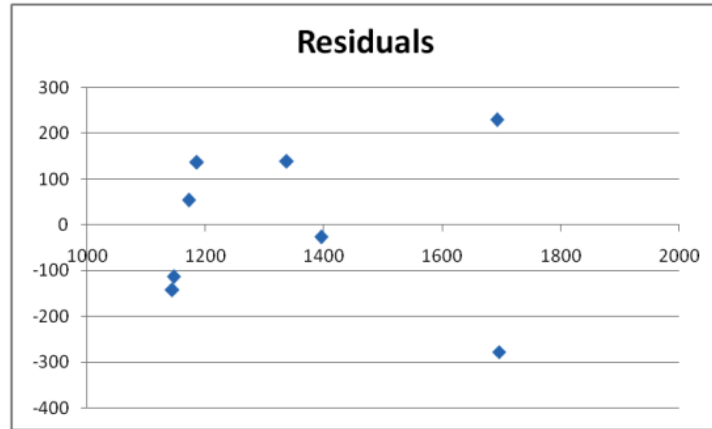
Figure 5.19: A residual plot for the regression model

Amplitude data is readily available for any position in the study site from GPS data on the internet. This could help predict rainfall data which was not readily available for the study sites.

5.13   Simulating weather data

Where temperature data is available, the altitude can be precisely predicted by the following model. Mean daily temperature was varied between 15°C and 25°C using the regression equation of altitude (alt) against average daily temperature (adt)

$$\text{Alt} = -155.26616(\text{adt}) + 4,667.36271$$

With $R^2 = 0.99574$ and standard error 29.048677 the corresponding altitudes were simulated. This equation was derived from actual data in eight places within the study region. Regression diagnostics for the simulated data were as follows:

Table 5.35: Regression coefficients for the simple linear regression model

Regression Statistics	
Multiple R	0.997286242
R Square	0.994579848
Adjusted R Square	0.994552474
Standard Error	33.40475278
Observations	200

Table 5.36: ANOVA table for average daily temperature on altitude

	df	SS	MS	F	Signif F
Regression	1	40542445.52	40542445.52	36332.34402	2.6E-226
Residual	198	220943.7467	1115.877509		
Total	199	40763389.27			

Table 5.37: Regression coefficients for average daily temperature on altitude

	Coefficients	Standard Error	t Stat	P-value
Intercept	4640.6658	16.6425	278.8442	6.8E-259
Adt	-153.8940	0.8074	-190.6105	2.6385E-226

5.14 Comparing climate data with bamboo growth measures

Having already determined that growth measures and climate data separately had strong linear relationships, we sought to compare one of the growth measures, Average number of culms per clump (Anoc), with rainfall, temperature and altitude. The multiple regression equation for average number of culms for *bambusa vulgaris* against the weather parameters was determined as follows.

Table 5.38: Regression statistics for average daily temperature, rainfall and altitude on Anoc

Regression Statistics	
Multiple R	0.994641348
R Square	0.989311412
Adjusted R Square	0.989147811
Standard Error	1.33758965
Observations	200

Table 5.39: ANOVA table for average daily temperature, rainfall and altitude on Anoc

	df	SS	MS	F	Significance F
Regression	3	32457.4619	10819.15397	6047.104894	7.6139E-193
Residual	196	350.6726301	1.789146072		
Total	199	32808.13453			

Table 5.40: Regression coefficients for mean daily temperature, rainfall and altitude on Anoc

	Coefficients	Standard Error	t Stat	p-value
Intercept	-12.05416786	14.24462613	-0.84622564	0.398458845
Average annual rainfall (mm)	-0.000219018	0.000713883	-0.306798329	0.759322571
Average daily temperature °C	-0.315294184	0.472982384	-0.666608724	0.505806017
Average altitude (M)	0.027872102	0.002983534	9.341974843	2.14383E-17

The predictors altitude and temperature were strongly correlated ($r = 0.99964$). Further, the coefficients are insignificant except altitude. This is a pointer to the problem of multicollinearity. To curb the problem, one of the two highly correlated independent variables was dropped since both predicted average number of culms (Anoc) with similar strength. Selecting altitude, we had a 2 independent variable regression model:

Table 5.41: Regression statistics for average rainfall and altitude on Anoc

Regression Statistics	
Multiple R	0.994629
R Square	0.989287
Adjusted R Square	0.989178
Standard Error	1.335702
Observations	200

Table 5.42: ANOVA table for average rainfall and altitude on Anoc

	df	SS	MS	F	Significance F
Regression	2	32456.67	16228.33	9096.09	8.82E-195
Residual	197	351.4677	1.7841		
Total	199	32808.13			

Table 5.43: Regression coefficients for average rainfall and altitude on Anoc

	Coefficients	Standard Error	t Stat	P-value
Intercept	-21.5464	0.378205	-56.9702	2.4E-124
Average yearly rainfall (mm)	-0.00012	0.000698	-0.17722	0.859518
Average altitude (M)	0.029815	0.000634	47.01264	5E-109

A hypothesis test on the coefficients was done:

$$H_0 : \beta_i = 0$$

against the alternative

$$H_1 : \beta_i \neq 0, \quad i = 0, 1, 2.$$

giving the following results

Table 5.44: Significance of regression coefficients for average rainfall and altitude on Anoc

β_i	Coefficients	t Stat	P-value	Decision
β_0	-21.5464	-56.9702	2.4E-124	Reject Ho
β_1	-0.00012	-0.17722	0.859518	Fail to reject Ho
β_2	0.029815	47.01264	5E-109	Reject Ho

The results implied that rainfall was not a significant predictor in the presence of Altitude. This means that a simple linear regression model of altitude on average number of culms was sufficient. Similar results were achieved when culm diameter, clump circumference and for both species of bamboo was predicted by the climate variables. Use of altitude was preferred since it is constant and is readily available for each point on the planet at the comfort of a study table.

5.15 Some regression models for predicting yield

5.15.1 Linear regression model

A simple linear regression model of the form

$$y = \beta_0 + \beta_1 t + \varepsilon$$

where y is the yield (number of culms per clump), β_0 is the yield at time $t = 0$, β_1 is the proportionate change in yield per unit increase in time t (months). ε is the random error term. Average number of culms per site was computed and recorded at various times of monitoring. The averages were used to fit a linear regression line. Figure 5.20 shows the fitted line. The simple linear regression model was

$$y = 0 + 0.4735t \tag{5.3}$$

The coefficient of determination R^2 for the regression model was $R^2 = 0.8374$. There was a strong linear relationship between number of culms and age from planting. Obviously, on extrapolation, the model will predict inconsistent values. For example, after 5 years (60 months) the model predicts 28.41 culms per clump. We expect more with time, without a maximum.
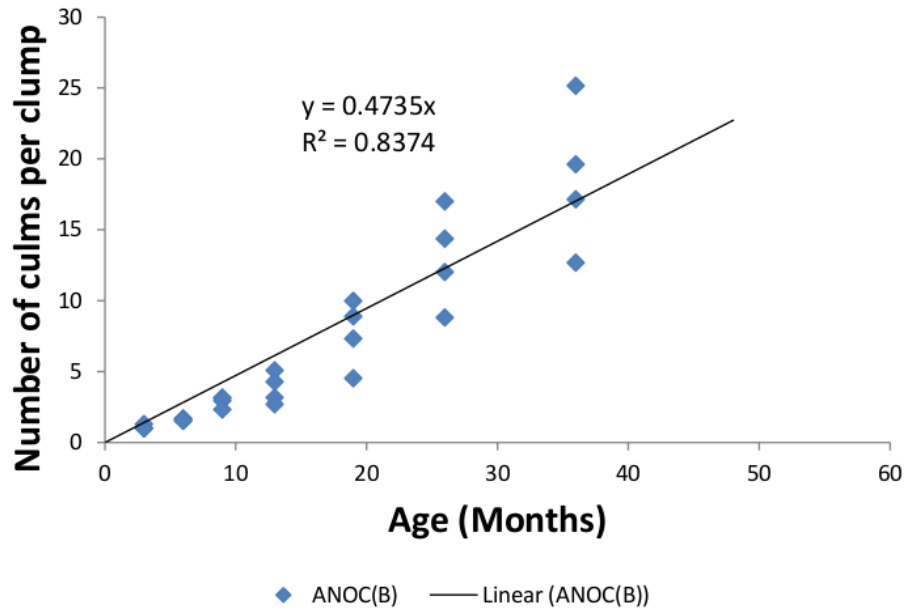
Figure 5.20: Average Number of Culms (*b. vulgaris*) linear model

5.15.2 Exponential fit

The exponential model is of the form

$$y = \beta_0 e^{\beta_1 t}$$

Here y and β_0 have similar meaning as in the linear case, β_1 has an exponential relationship with yield per unit time t (months). The exponential model is plotted in Figure 5.21. The model equation was:

$$y = 1.0925 e^{0.0867t} \tag{5.4}$$

The coefficient of determination R^2 for the exponential model was $R^2 = 0.9059$. This was higher than the linear model case. On extrapolation however, the model will predict worse off than the linear counterpart. For example, after 5 years the model predicts 198.44 culms.
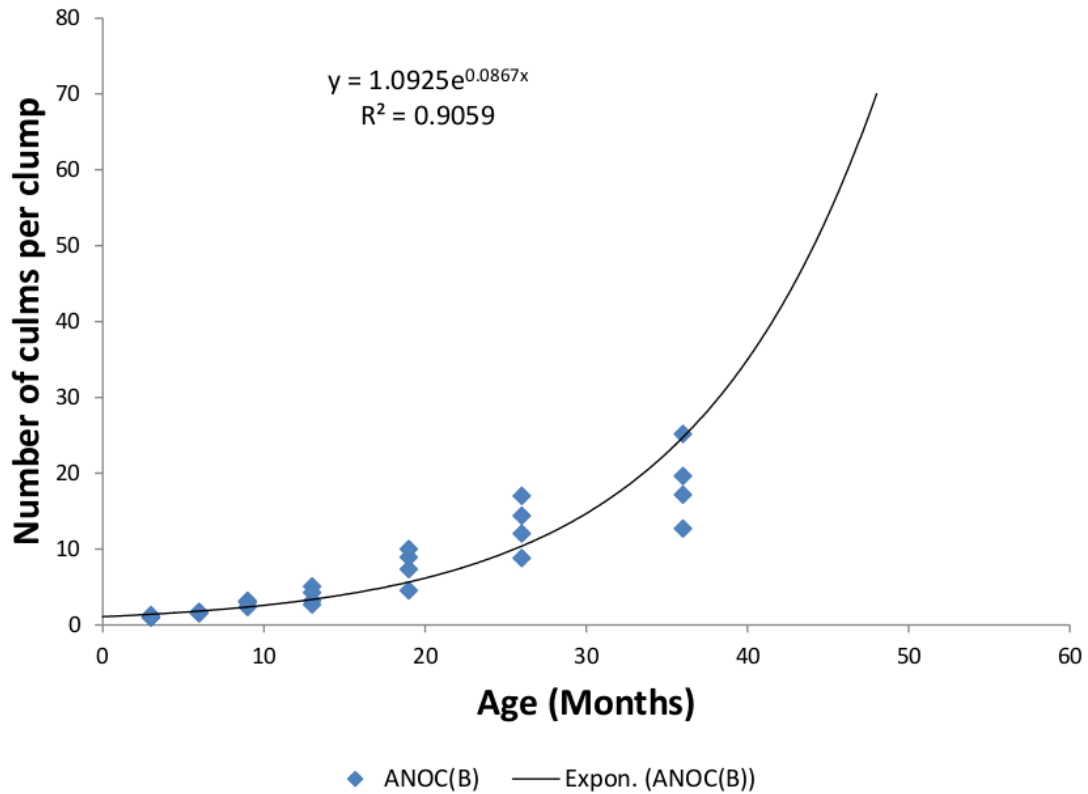
Figure 5.21: Average Number of Culms (*b. vulgaris*) exponential model

5.15.3 The power function

The power model is of the form

$$y = \beta_0 t^{\beta_1}$$

Here when $t = 0, y = 0$ regardless the value of β_0 or β_1. The power function is plotted in Figure 5.22. The model equation (equation 5.5) was:

$$y = 0.2232t^{1.1941} \tag{5.5}$$
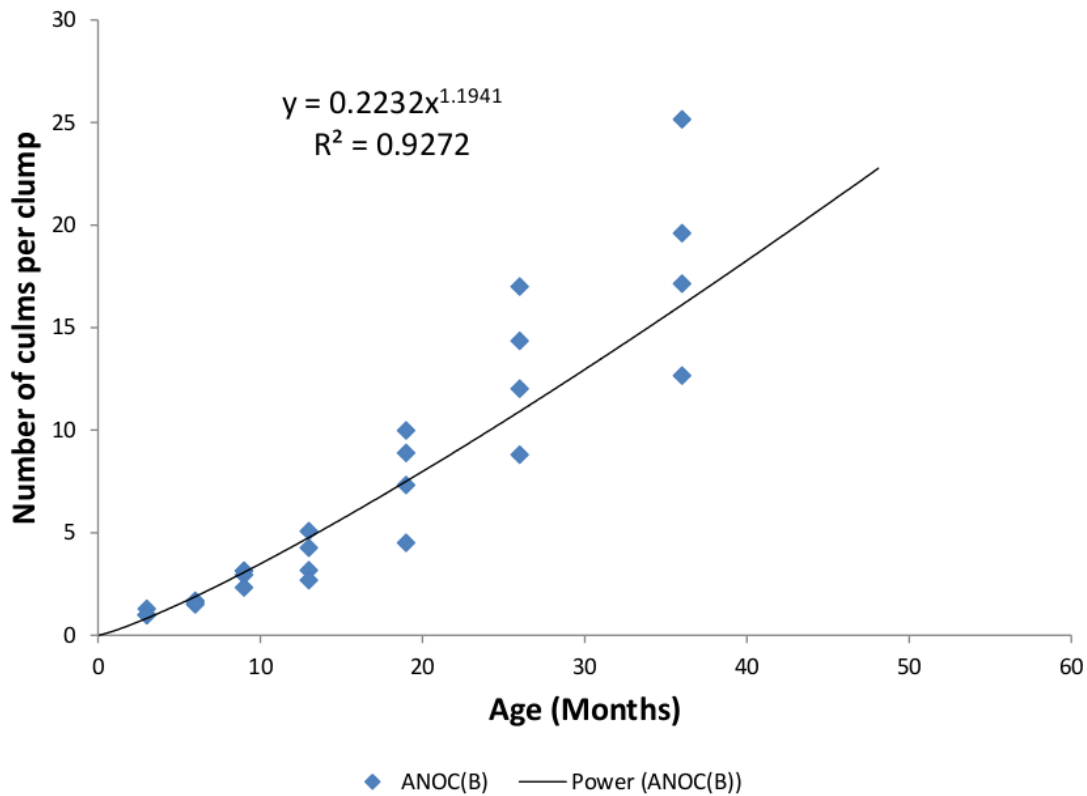
Figure 5.22: Average Number of Culms (*b. vulgaris*) power model

The coefficient of determination R^2 for the power model was $R^2 = 0.9272$. This is higher than both the linear and exponential model cases. On extrapolation, say, after 5 years the model predicts 29.65 culms. This is a more reasonable estimate than that of the exponential model.

5.15.4 The quadratic function (polynomial of degree 2)

A polynomial regression model is of the form

$$y = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon$$

Figure 5.23 shows the fitted curve. The regression model was

$$y = -0.3622 + 0.3051t + 0.0065t^2 \tag{5.6}$$

The coefficient of determination R^2 for the regression model was $R^2 = 0.876$. On extrapolation, the model predicted higher values. For example, after 60 months the model yielded 43.14 culms per clump.
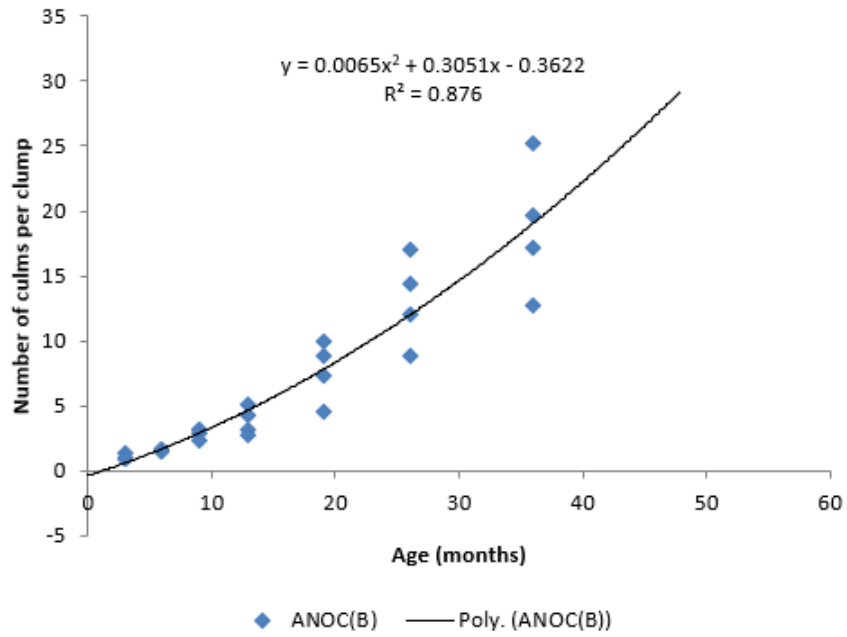
Figure 5.23: Average Number of Culms (*b. vulgaris*) quadratic model

5.15.5 The cubic function (polynomial of degree 3)

A cubic regression model is of the form

$$y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \varepsilon$$

Figure 5.24 shows the fitted curve. The regression model was

$$y = 1.5975 - 0.2333t + 0.0409t^2 - 0.0006t^3 \qquad (5.7)$$
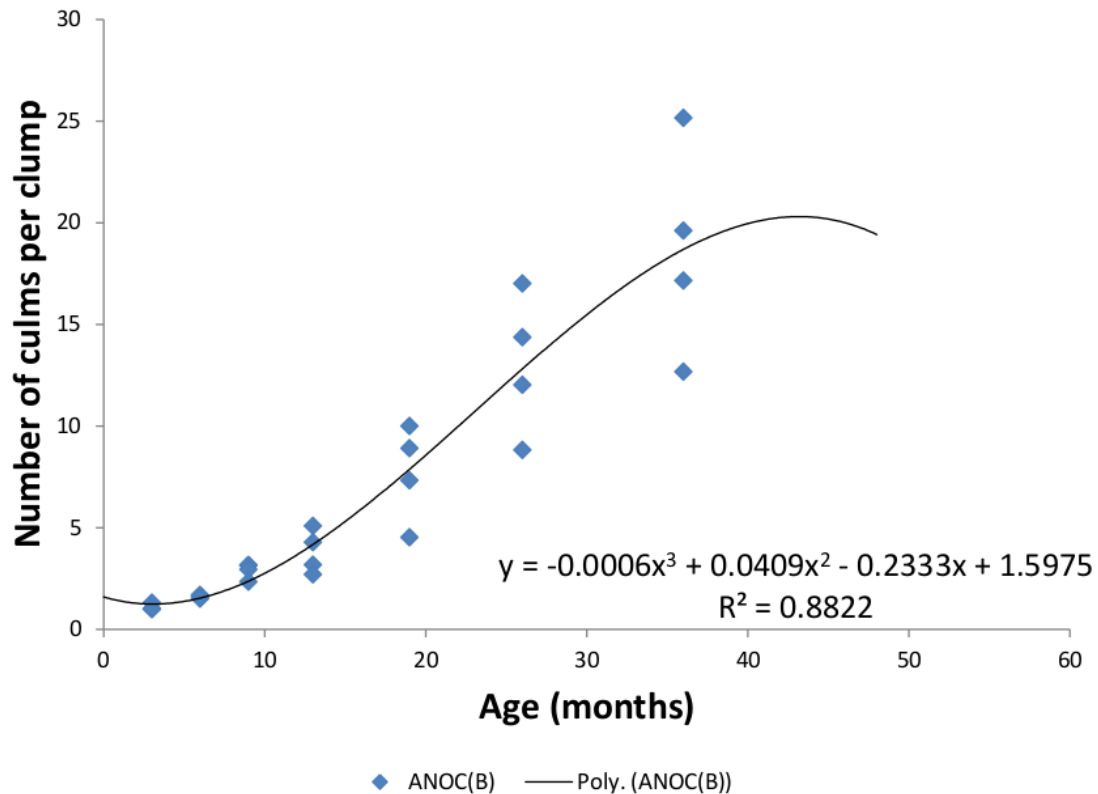
Figure 5.24: Average Number of Culms (*b. vulgaris*) cubic model

The coefficient of determination R^2 for the regression model was $R^2 = 0.8822$. On extrapolation, the model predicted inconsistent values. For example, after 5 years (60 months) the model yielded 5.24 culms per clump. This is an under estimate.

5.15.6   The 5-parameter logistic model

Having determined the number of culms per clump as a measure of productivity, there was need to determine the best fit model which not only predicts expected number of culms within the interval of data, but also can be used to forecast future values past the three years of available data. The 5-parameter logistic (Richard's) model was suggested.

$$P(t) = B + \frac{T - B}{[1 + 10^{b(t_{mid} - t)}]^s}$$

where $P(t)$ is the population at time t, B is the initial population, T is the carrying capacity, b and s are constants.

Simulating the 5-Parameter logistic model

To determine the line of best fit, simulations of various slopes for the model were made

95

and results plotted (see Figure 5.25).
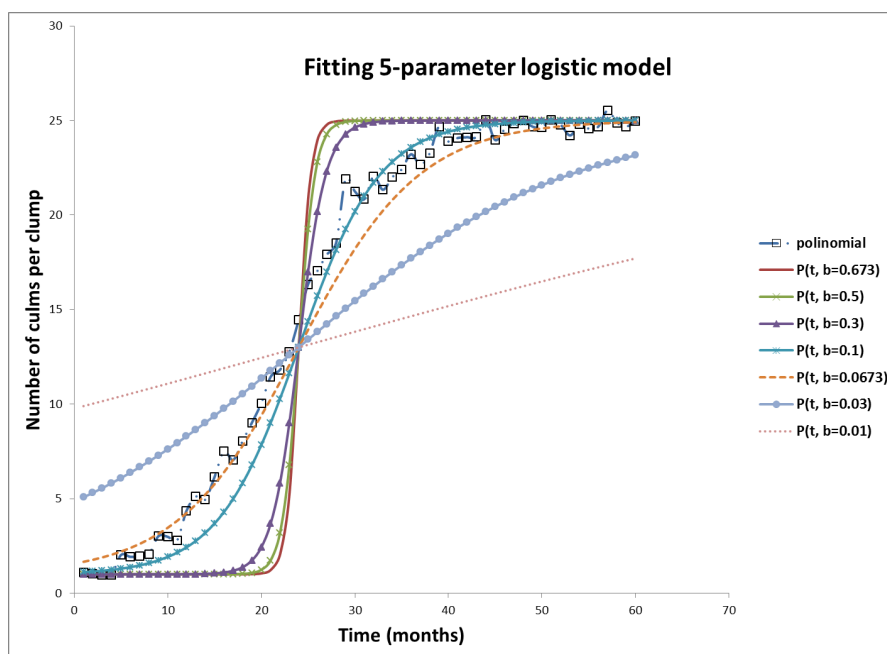
Figure 5.25: Fitting a 5-parameter logistic model

Table 5.45 shows various values of the hill slope b and corresponding correlation of the resultant logistic curve with the polynomial curve. The bold type coefficients were highest, giving a hint on the appropriate value of b.

Table 5.45: Parameter estimation for logistic model

	Polyno mial	P(t, b=0.673)	P(t, b=0.5)	P(t, b=0.3)	P(t, b=0.1)	P(t, b=0.06 73)	P(t, b=0.03)	P(t, b=0.01)
Polynomial	1.0000							
P(t, b=0.673)	0.9198	1.0000						
P(t, b=0.5)	0.9256	0.9995	1.0000					
P(t, b=0.3)	0.9388	0.9954	0.9978	1.0000				
P(t, b=0.1)	**0.9613**	0.9641	0.9696	0.9811	1.0000			
P(t, b=0.0673)	**0.9663**	0.9444	0.9503	0.9636	0.9961	1.0000		
P(t, b=0.03)	0.9176	0.9005	0.9066	0.9213	0.9700	0.9866	1.0000	
P(t, b=0.01)	0.8744	0.8695	0.8755	0.8902	0.9441	0.9666	0.9950	1.0000

The value of the gradient that could make the logistic model run closest to the polynomial fit was $P(t, b = 0.0673)$. The correlation between the two curves was 0.9663 which is highest among the rest. To ascertain the graphical results, R-software [70] was used to simulate data for the 5-parameter logistic model. The unknown parameters were determined as in equation (5.8). $s = 1$ implies that the model is symmetric at t_{mid}.

$$P(60) = 1 + \frac{25 - 1}{[1 + 10^{.0673(24-60)}]^1} \tag{5.8}$$

These parameters were similar to those estimated in Figure 5.26. The goodness of fit statistic for the regression model was GOF=0.9967. On extrapolation, the model predicted consistent values. For example, after 60 months the model yielded 25 culms per clump. Thus depending on the maximum number T, the model will never give higher values than T. The cumulative density function for the number of culms is plotted in Figure 5.25.
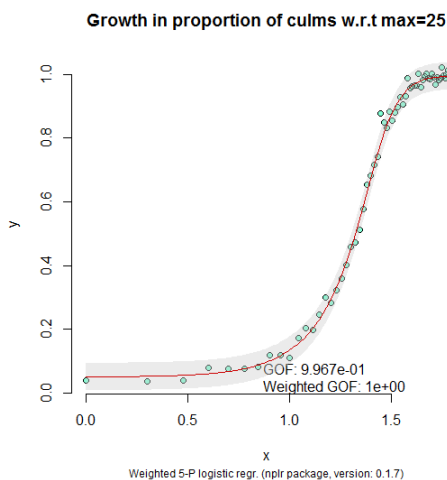
Figure 5.26: Cumulative density of culms (y) in log time (x) for *b. vulgaris*

Further a bootstrap Monte Carlo simulation was carried out with interesting results. We were able to fit the bamboo growth in number of culms to a Gamma density with parameters $G(X, \alpha, \beta) = G(X, 4, 3)$. Figure 5.27 shows the actual data (in dots) and Gamma distribution data (in small squares).
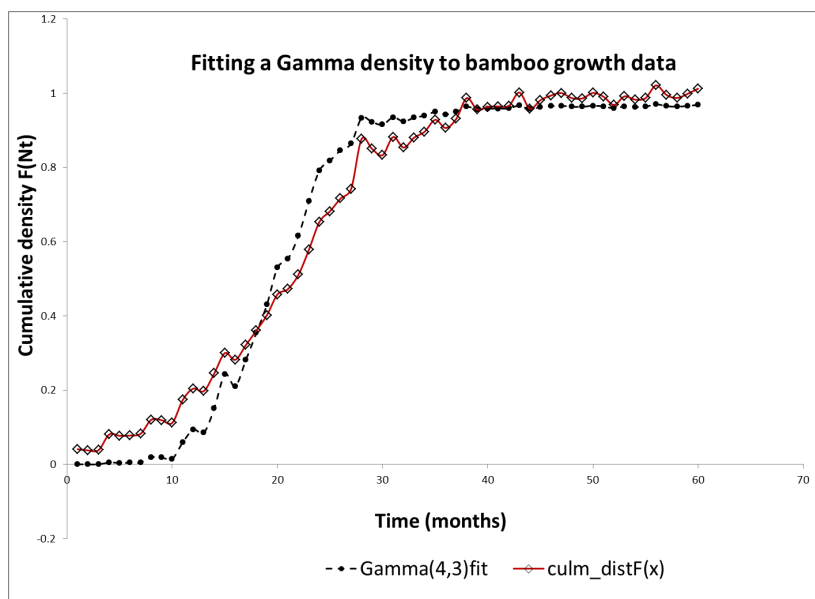
Figure 5.27: Gamma (4, 3) simulated fit for bootstrap Monte Carlo number of culms

Further, a two-sample Kolmogorov-Smirnov test gave the results in Table 5.46. Codes for the tests is appended (see appendix A.3, A.4 and A.5).

Table 5.46: Kolmogorov-Smirnoff GOF test

b	0.673	0.5	0.3	0.1	0.0673	0.03	0.01
D	0.4667	0.45	0.4000	0.15	0.1000	0.4000	0.5667
P-value	4.2x10^{-6}	1.058x10^{-5}	0.0001	0.5095	0.9251	0.0001	8x10^{-9}

The Kolmogorov D statistic is the absolute max distance (supremum) between the CDFs of the two samples. The closer this number is to 0 the more likely it is that the two samples were drawn from the same distribution. The hypotheses read:

H_0 The two samples have the same distribution.

H_1 The two samples have different distributions.

The p-value returned by the $K - S$ test has the same interpretation as other p-values. We reject the null hypothesis that the two samples were drawn from the same distribution if the p-value is less than the set significance level. From Table 5.46, the lowest $D = 0.1000$ is at $b = 0.0673$, the p-value 0.9251 suggests that we have no sufficient reason to reject the null hypothesis. Thus we conclude that the selected b gives the best fit model for the 5 parameter logistic equation.

Chi square goodness of fit test also showed that the best fit value for b was 0.0673. The p-value= 0.01815 which is less than 0.05. Thus there was no significant difference between actual and simulated data.

The full model with its parameters is

$$P(t) = 1 + \frac{25 - 1}{\left[1 + 10^{.0673(24-t)}\right]^1}$$

5.16 Fitted models and associated population predictions

In chapter 4, a number of models were discussed. We started with the classical logistic model by Verhulst, which has over the years become a subject of many modifications and applications. We also looked at the Beverton - Holt discrete logistic model and suggested modifications to allow for a delay to onset of harvest (equation 4.7). We also suggested a stochastic version of the model (equation 4.8). A logistic model with

constant harvest (equation 4.17) and with proportional to size harvest (equation 4.21) were characterized and the optimal harvesting strategy for bamboo found to be 9.8 culms per clump per year. Further, Stochastic logistic harvest equations with both harvest and growth random (equation 4.58, 4.70, 4.80) were formulated and analyzed. Simulated trajectories were also used to analyze their behaviors. It was on average determined that a stochastic system will behave in a similar manner as its deterministic counterpart as long as noise was kept sufficiently low. The time scale for stochastic modeling was in months while that of deterministic models was in years. That variation made us determine that harvesting in shorter, almost continuous intervals will in the long ran yield more bamboo than wait for a year to cut culms. While yearly deterministic harvesting led to averagely 9.8 culms per clump as yield, monthly harvesting led to a sustainable optimum of 15.5 culms per clump each year. Other models that were fitted in chapter 5 are summarized in the following table.

Table 5.47: Summary of fitted models

Model	Number of culms at t=36	Number of culms at t=36	Goodness of fit
Linear $y(t) = 0.4735t$	17.05	28.41	$R^2 = 0.8374$
Exponential $y(t) = 1.0925e^{0.0867t}$	24.77	198.44	$R^2 = 0.9059$
Power $y(t) = 0.2232t^{1.1941}$	16.11	29.65	$R^2 = 0.9272$
Polynomial order 2 $y(t) = 0.0065t^2 + 0.3051t - 0.3622$	19.05	43.14	$R^2 = 0.8760$
Polynomial order 3 $y(t) = -0.0005t^3 + 0.0367t^2 - 0.1751t + 1.5387$	19.47	15.15	$R^2 = 0.8822$
5- parameter logistic $P(t) = 1 + \dfrac{25 - 1}{[1 + 10^{.0673(24-t)}]^1}$	20.47	23.88	GOF = 0.9991
Gamma(4,3), $F(t, \alpha, \beta)$	24.94	25.00	Signif. 0.018

The main growth measures that determine bamboo harvest volume were studied. These include survival rate, average height, and average culm diameter and average number of culms per clump. At three years from planting the population of culms was studied with the following results achieved.

Table 5.48: A summary of major parameter measures achieved during the research

Site	% survival		Mean height at 36 months		Mean culm diameter at 36 months		Mean number of culms per clump at 36 months xx.xx (95% CI)	
	B. vulgaris	D. giganteus	B. vulgaris	D. giganteus	B. vulgaris	D. giganteus	B. vulgaris	D. giganteus
Rangwe	95%	80%	7.26	9.94	4.67	7.89	16.62 (1.71)	12.36 (1.23)
Suna E.	82%	58%	6.32	7.64	4.75	6.83	23.05 (3.77)	22.35 (5.37)
Kuria	58%	22%	6.79	8.93	4.87	7.77	18.33 (1.69)	16.55 (1.87)
Suba	90%	60%	5.85	7.41	4.62	6.78	10.17 (4.52)	11.67 (1.94)

5.17   Discussion

5.17.1   Bamboo growth parameters

The study involved determining various characteristics of two species of bamboo namely, *bambusa vulgaris* and *dendrocalamus giganteus*. This were planted in randomly distributed farms Suna East and Kuria West sub-counties of Migori County and Suba together with Rangwe sub-counties of Homabay County in Kenya. The growth parameters that were measured include survival proportions. It was found that *b. vulgaris* had a higher probability of survival than *d. giganteus* in all the study sites. It was also noted that bamboo survived more in Homabay County than in Migori (see figure 5.1).

Culm height was also studied. It was demonstrated from data that *d. giganteus* achieved higher heights at all times from planting to maturity than *b. vulgaris* culms. At the age of 3 years, *b. vulgaris* achieved heights up to 6.5 m while *d. giganteus* averaged 8.5m (see figure 5.4). Clump diameter was also measured where *b. vulgaris* was found to spread out more rapidly, hence achieving larger clump diameters in comparison to the tightly clumping *d. giganteus*. However, it was noted that mean crown cover diameters of both species was not significantly different.

Culm diameter was also measured and analyzed. This was done at breast height (dbh). It was found that *d. giganteus* had larger culm diameter than that of *b. vulgaris* on average. By the age of 3 years, *b. vulgaris* diameter was ranging between 4.6 and 4.9 cm while *d. giganteus* attained between 6.7 and 7.9 cm diameter. This variation in culm diameter is also evident in Figure 5.7. Height and culm diameters were highly comparable with a correlation coefficient of at least 0.97 in all sites for both species.

The most important measure that was studied at length was the number of culms in each

clump. This has a direct proportionality with the amount of yield of a farm. Figure 5.10 shows a comparison of reproduction of bamboo culms among the 4 study sites. On average, Kuria West registered a higher culm population growth rate than the rest of the sites. Suba was worst performing. Among the two species, at early age (up to 2 years), the clump population was similar (see Figure 5.12). The variation came in towards the third year where *b. vulgaris* registered averagely higher culm counts than *d. giganteus*. The number of culms per clump was found to be estimated approximately by the clump circumference. A linear model was suggested (see figure 5.13). A one way ANOVA test (Table 5.24) showed that there was significant difference performance among the four study sites. Table 5.26 summarizes the 95% confidence intervals for number of culms per clump for the two species in all sites.

Weather data of the four study sites was acquired from climate-data.org freely available online. This was used to determine relationships between weather and the productivity of sites. A relationship between altitude and mean daily temperature was established. It was a strong linear fit with the coefficient of determination 0.9948. Annual average rainfall was also compared with altitude in a multiple linear regression model where it was found to have an insignificant coefficient. Altitude was therefore selected as a representative variable for climate.

5.17.2   Determining most suitable growth model

A number of culms per clump was fitted with various regression curves including linear, exponential, power, curvilinear (polynomial) and a 5-parameter logistic (Richard's) model. Only the logistic model was found not only to fit the collected data well but maintained a reasonable path on extrapolation to forecast future population values. The model had a low variance inflation factor (VIF) and a coefficient of determination of 0.999. It was from this finding that further studies on logistic harvest modeling was required.

5.17.3   Bamboo growth and harvest models

In chapter 4, we considered various models, all of which were variants of the Verhulst logistic growth model. The Beverton - Holt discrete logistic model was discussed in section 4.3 where bamboo data were fitted to the model with interesting results (see Figure 4.1(a), (b), and 4.2(a), and (b). Our contribution in this model was the suggestion to have no harvest in the first two years of growth and start harvesting at the end of the third year.

We further suggested a Stochastic Beverton-Holt model of the form (equation (4.8)). This was simulated 100 times to produce the trajectory (Figure 4.2(b). The noise was assumed to be uniformly distributed between $[-\sigma, \sigma]$ where σ is the noise intensity. In the same model, with the requirement that only mature culms (3 years old) are to be harvested, the harvest $h = N_{(t-2)} - N_{(t-3)}$ was considered. For the non-volatile case, an equilibrium average harvest of 12.5 culms per clump was suggested (see Figure 4.2(a). It was noted that, both noise and harvesting can lead to population extinction in finite time.

In section 4.5, a logistic growth model with constant harvesting rate H was analyzed (equation (4.9)). It was determined that the critical harvest that could keep the system at equilibrium with maximum returns is the lower neighborhood of $H_c = \frac{rK}{4}$. An explicit solution for the system (see equation (4.17)) was derived. Further, it was shown that as long as the optimum harvesting level was not passed, the system remained persistent. Higher than optimum harvest was shown to cause extinction in finite time (see Figure 4.3). Here again a delayed onset of harvest was advised (equation (4.20)). The maximum critical harvest was found to be 9.8 culms per clump annually (see equation (4.21)).

We also investigated a deterministic logistic model (equation (4.23)) with harvest proportional to population size. Its explicit solution was derived (equation (4.31)) the growth rate for bamboo was computed (equation (4.32)) and found to be 1.56868. As was the case in previous models, a delayed onset of harvesting was suggested (equation (4.33)). Harvest at equilibrium was found to be 9.8 culms per clump each year, as was the constant harvest case. Finally, a stochastic logistic model was formulated. We considered a case where the growth and harvest rates are both volatile. The yield density was computed using the Fokker-Planck equation. Further, the mean and variance of yield were computed. The resultant model was tested using actual bamboo data at various levels of noise. It was concluded that the average behavior (drift) of the process was the same as the deterministic case. As long as the noise levels were kept low, the system on average was shown to be persistent within 0 and K. Assuming that harvest was systematic (with a consistent rate of cutting each time), a single noise system was modelled (equation (4.77) to (4.80)). No clear stable state was achievable due to the continuous stochastic system.

5.17.4 Optimal bamboo harvesting strategy

According to Arori et. al [9] an acre of land can contain up to 220 clumps of *bambusa vulgaris* spaced 5 square meters apart. Considering that approximately 80% of the seedlings will grow to mature clumps, its expected that 176 clumps will be available. During this research, we deduced that an equilibrium harvest strategy is to cut 9.8 culms per clump

per year. This yields a yearly harvest of 1725 culms per acre. Suppose a market price of Kshs. 100 per culm [56], an annual income of Kshs.172,480 will be realized. This translates to a monthly income of Kshs.14, 373.33. This is a reasonable amount considering that very little effort is needed to maintain clumps once they mature. Besides this, if one opts to produce bamboo items such as furniture, the income becomes much higher.

CHAPTER SIX

SUMMARY, CONCLUSION AND RECOMMENDATION

6.1 Introduction

In this chapter we have a summary of how the study was executed and the accomplishments we were able to make. We draw conclusions from the results in relations to the stated objectives. Further, limitations and recommendations for use and for further study are also given.

6.2 Summary

The study involved bamboo growth and harvesting data collection and analysis. Mathematical population models suitable for bamboo plantations were derived. During modeling both Ordinary Differential Equations (ODE) and Stochastic Differential Equations were used. The logistic population growth model was the basis of all derived models in order to understand bamboo growth behavior and characterize parameters such as number of culms per diameter and clump circumference, time series data that was taken during the experimental part of the study were analyzed.

Model fitting,where linear and non-linear models were tested for goodness of fit, was used to determine the best model. A 5-parameter-logistic regression model was found to fit actual data best.

Due to the finding, the logistic model was studied further by exploring equilibrium behavior and by involving a harvesting component. Both constant harvest rate and harvest proportional to size was considered.

Finally, since Bamboo populations are prone to environmental disturbance, stochastic logistic population models were formulated and analyzed.

Optimum harvest under deterministic conditions was found to be constant at 10 culms per clump. Under stochastic conditions the optimum harvest could vary depending on the volatility in growth and harvesting.

6.3 Conclusion

In relation to the set objectives, the study was successful. From experimental data, at the point of harvesting Bamboo, 36 months after planting, a clump of *bambusa vulgaris* had averagely a population of 18 culms while that of *dendrocalamus giganteus* had 15 culms. The average height of a b. vulgaris culms was 6.55M while that of d.giganteus was 8.68M.The culm diameter was 8.68M. The culm diameter was 4.73cm and 7.31cm respectively. Various regression models were fitted and the best found to be a 5-parameter Logistic model with initial population size 1 culm, carrying capacity 25 culms and an estimated population 20 culms at age 36 months and 24 culms at age 60 months.

A logistic regression with constant harvest and with harvest proportional to population size was formulated. In both cases, harvesting 10 culms per clump per year was found optimum.

Finally, a stochastic logistic model with Brownian growth and harvest was formulated. The general trend in population growth and harvesting was similar to that of non-stochastic case. It was concluded that low noise resulted in a stable system while high noise led to extinction.

6.4 Limitations and Recommendations

Situations under which bamboo could go extinct were analysed in each model presented. Mainly, it was noted that without noise the harvest quantity could determine persistence of the population or extinction in case some critical level (more than 9.8 culms per clump) was surpassed. Later in the stochastic model, Noise levels teamed up as major threats to population survival. The higher the noise, the higher the probability that the system could go extinct. Suppose all culms are cut, the models will perceive the scenario as extinction. However, for bamboo, this is not the case. New culms will shoot up rampantly and in two to three years, another harvest will be available. How to represent this scenario will be a good ground for further model modification.

From the reported results, the research was largely successful in meeting the set objectives.

This was not without challenges though. During the research process retagging was necessitated by loss of tags from the plants. Cattle and other domestic animals were reported to have grazed on bamboo leaves hence either causing the death of clumps or retarding growth. Poor farm management was also a serious problem.

At the modeling level, we recommend that further studies in this regime of mathematical ecology be made to determine the effect of noise on the carrying capacity of bamboo. Further the effects of watering and adding manure or fertilizer in clumps on the carrying capacity is a good extension to this work.

In the course of our literature review, Anderson et al [8] noted that a logistic equation with a simple stochastic carrying capacity has not been successfully solved. This is another area for further research.

REFERENCES

[1] Addison A., Dale G., and Baker S. (2010). Estimating Crown and Understory Biomass in Operational Harvests. Technical report, in Proc. 33rd annual meeting of the council on Forest Engineering, D. Mitchell and T. Gallagher eds. Center for Forest Business, University of Georgia, Athens, GA 30602-2152.

[2] Al Sharawi Z. and Rhouma M. B. H. (2009). The Beverton-Holt model with periodic and conditional harvesting. *Journal of Biological Dynamics*, 3(5):463–478.

[3] Allen L. J. S. (2017). A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infectious Disease Modeling*, 2:128–142.

[4] Alvares L. H. R. and Shepp L. A. (1998). Optimal harvesting of stochastically fluctuating populations. *J. Math Biol.*, 37:155–177.

[5] Ambika K. and Rajagopal B. (2017). In vitrio antimicrobial and antiproliferative activity of bambusa vulgaris. *Ijppr.Human*, 9(1):10–22.

[6] American Bamboo Society (2006). Bamboo Species Source List No. 26. `http://www.abssocal.org/pdf/BambooSourceList2006/2006SourceListIntro.pdf`. Accessed: 2009-11-18.

[7] American Bamboo Society (2008). Bamboo Species source list no. 28. `http://www.bamboo.org/SpeciesSourceList.html`. Accessed: 2009-11-18.

[8] Anderson C., Jovanoski Z., Sidhu H. S., and Towers I. N. (2016). Logistic equation is a simple stochastic carrying capacity. *ANZIAM J.*, 56:C431–C455.

[9] Arori W., Kibwage J. K., Netondo G. W., and Onyango F. (2013). A Markov Model for Bamboo Harvest Forecasting in South Nyanza Region, Kenya. *Advanced Journal of Agricultural Research*, 1(4):45–50.

[10] Avsar M. D. (2004). The relationships between diameter at breast height, tree height and crown diameter in calabrian pines (*Pinus brutia Ten.*) of Baskonus Mountain, Kahramanmaras, Turkey. *J. Biol. Sci.*, 4:437–440.

[11] B. Keating A., P. Carberry S. G. Hammer M. P., Holzworth J. R. N., and Huth N. (2003). An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, 18:268–288.

[12] Bechtold W. A. (2003). Crown-diameter predictions models for 87 species of stand-grown trees in the Eastern United states. *South. J. Appl. For.*, 27:269–278.

[13] Bechtold W. A. (2004). Largest-crown-width prediction models for 53 species in the western united states. *West. J. Appl. For.*, 19:245–251.

[14] Beddington J. R. and M. R. (1977). May, Harvesting natural population in a randomly fluctuating environment. *Science*, 197:463–465.

[15] Benitez J. Y. and M. Rivero Vidal A R. J. . R. (2003). Estimación del diámetro de copa a partir del diámetro normal (d1,3) en plantaciones de casuarina equisetifolia. *Forst. Invest. Agrar: Sist. Recur. For.*, 12:37–41.

[16] Beverton R. J. H. and Holt S. J. (2004). *On the Dynamics of Exploited Fish Populations.* Blackburn Press, New Jersey.

[17] Biging G. S. and Dobbertin M. (1992). A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. *For. Sci.*, 38:695–720.

[18] Birch C. P. D. (1999). A new generalized logistic sigmoid growth equation compared with the Richards growth equation. *Ann Bot*, 83:713–723.

[19] Bohner M. and Warth H. (2007). The Beverton-Holt dynamic equation. *Applicable Analysis*, 86(8):1007–1015.

[20] Brag D. C. (2001). A local basal area adjustment for crown width prediction. *North. J. Appl. For.*, 18:22–28.

[21] Brauer F. and Sanchez D. A. (1975). Constant rate population harvesting: equilibrium and stability. *Theoret. Popul. Biol.*, 8:12–30.

[22] Brauer F. and Sanchez D. A. (2003). Periodic environments and periodic harvesting. *Natur. Resource Modell.*, 16:233–244.

[23] Cañadas N. (2000). *Pinus pinea L. en el Sistema Central (Valles del Tiétar y del Alberche): desarrollo de un modelo de crecimiento producción de piña.* PhD thesis, E.T.S.I. de Montes, Universidad Politcnica de Madrid.

[24] Cam E., Link W. A., and E.G C. (2002). Individual covariation in lifehistory traits: Seeing the trees despite the forest. *Am Nat*, 159:96–105.

[25] Clark C. W. (1990). *Mathematical Bioeconomics: the Optimal Management of Renewable Resources*. JohnWiley and Sons, NewYork.

[26] Clark J. S., LaDeau S., and Ibanez I. (2004). Fecundity of trees and the colonization - competition hypothesis. *Ecol Monogr*, 74:415–442.

[27] Climate-data.org (2017a). Climate: Homa Bay (Rangwe). `http://en.climate-data.org/region/2587/`. Accessed: 22-01-2017.

[28] Climate-data.org (2017b). Climate: Homa Bay (Suba-Sindo). `http://en.climate-data.org/location/103177`. Accessed: 22-01-2017.

[29] Climate-data.org (2017c). Climate: Migori (Kuria-West). `http://en.climate-data.org/location/104104/`. Accessed: 22-01-2017.

[30] Climate-data.org (2017d). Climate: Migori (Suna-East). `http://en.climate-data.org/location/103149/`. Accessed: 22-01-2017.

[31] Coleman B. D. (1979). Nonautonomous logistic equations as models of the adjustment of populations to environmental change. *Mathematical Biosciences*, 45(3-4):159 – 173.

[32] Damgaard C., Weiner J., and Nagashima H. (2002). Modelling individual growth and competition in plant populations: growth curves of Chenopodium album at two densities. *J Ecol*, 90:666–671.

[33] Daniels R. F., Burkhart H. E., and Clason T. R. (1986). A comparison of competition measures for predicting growth of loblolly pine trees. *Can. J. For. Res.*, 16:1230–1237.

[34] Diver S. (2001). Bamboo. Technical report, A Multi-purpose Agroforestry Crop. Current Topic. Appropriate Technology Transfer for Rural Areas (ATTRA).

[35] Doust R. M. and Saraj H. M. (2015). The logistic modeling population; having harvesting factor. *Yugoslav Journal of Operations Research*, 25(1):107–115.

[36] Espiloy Z. B. (1991). Effect of age on the physico-mechanical properties of some Philippine bamboo. Technical report, In Bamboo in Asia and the Pacific. Proceedings of the International Bamboo Workshop, Chiangmai, Thailand.

[37] Fan M. and Wang K. (1998). Optimal harvesting policy for single population with periodic coefficients. *Math. Biosci.*, 152:165–177.

[38] Fox J. C., Ades P. K., and Bi H. (2001). Stochastic structure and individual-tree growth models. *For. Ecol. Manage*, 154:261–276.

[39] Gill S. J., Biging G. S., and Murphy E. C. (2000). Modelling conifer tree crown radius and estimating canopy cover. *For. Ecol. Manage*, 126:405–416.

[40] González-Sánchez M., Cañellas I., and Montero G. (2007). Generalized height-diameter and crown diameter prediction models for cork oak forests in spain. *Invest. Agrar: Sist. Recur. For.*, 16:76–88.

[41] Grozdanovski T. and Shepherd J. J. (2007). Slow variation in the Gompertz model. *ANZIAM J.*, 47:C541–C554.

[42] Grozdanovski T., Shepherd J. J., and Stacey A. (2009). Multi-scaling analysis of a logistic model with slowly varying coefficients. *Applied Mathematics Letters*, 22:1091–1095.

[43] Grozdanovski T., Shepherd J. J., and Stacey A. (2010). Transitions in density dependent harvesting of a logistic population in a slowly varying environment. *ANZIAM J.*, 51:C393–C408.

[44] Hann D. W. (1997). Equations for predicting the largest crown width of stand-grown trees in Western Oregon. Technical report, Forestry Research Laboratory, Research Contribution 17, Oregon State University, Corvallis.

[45] Hann D. W. (1999). An adjustable predictor of crown profile for stand-grown Douglas-fir trees. *For. Sci.*, 45:217–225.

[46] Hunt R. (1982). Plant Growth Curves. Technical report, Baltimore, MD: University Park.

[47] Huth N. I., Carberry P. S., Brennan P. P. L., and Keating B. A. (2002). A framework for simulating agroforestry options for the low rainfall areas of Australia using APSIM. *European Journal of Agronomy*, 18:171–185.

[48] Islam A, A. Q. Miah A. H. and Rasul G. (2015). Growth and development of *bambusa vulgaris* schrad ex wendl. planted in the coastal homesteads of Bangladesh. *J Asiat. Soc. Bangladesh Sci*, 41(2):123–129.

[49] Jingliang L. and Wang K. (2015). Almost sure permanence of stochastic single species models. *J. Math. Anal. Appl.*, 422:675–683.

[50] Kibwage J. K., Netondo G. W., et al. (2008). Growth performance of bamboo in tobacco-growing regions of South Nyanza,kenya. *African Journal of Agricultural Research*, 3:714–722.

[51] Kigomo N. B. (2007). Guidelines for growing bamboo. Technical report, KEFRI Guideline Series: No. 4, Kenya Forestry Research Institute.

[52] Lewy P. and Nielsen A. (2003). Modelling stochastic fish stock dynamics using Markov Chain Monte Carlo ICES. *Journal of Marine Science*, 60.

[53] Li W. and Wang K. (2010). Optimal harvesting policy for general stochastic logistic population model. *J. Math. Anal. Appl.*, 368:420–428.

[54] Liu Meng W. K. (2012a). On a stochastic logistic equation with impulsive perturbations. *Computers and Mathematics with Applications*, 63:871–886.

[55] Liu Meng W. K. (2012b). Stationary distribution, ergodicity and extinction of a stochastic generalized logistic system. *Applied Mathematics Letters*, 25.

[56] Magati P. and J. Kibwage K, Seth Omondi G W. O. A. (2012). A Cost-benefit Analysis of Substituting Bamboo for Tobacco: A case study of smallholder tobacco farmers in South Nyanza, Kenya. *Science Journal of Agricultural Research & Management*, 2012:ISSN 2276–6375.

[57] Marshall D. D., Gregory P. J., and Hann D. W. (2003). Crown profile equations for stand-grown Western Hemlock trees in Northwestern Oregon. *Can. J. For. Res.*, 33:2059–2066.

[58] McPherson E. G., Peper J. R. S. . J., and Xiao Q. (1999). Benefit cost analysis of Modesto's municipal urban forest. *J. Arboric*, 25:235–248.

[59] Moore A. D., Huth D. H. N. H. N., and Robertson M. (2007). The common modelling protocol: A hierarchical framework for simulation of agricultural and environmental systems. *Agricultural Systems*, 95:37–48.

[60] Morteza K. and Kiaee N. (2014). Stochastic Dynamical Theta-Logistic Population Growth Model. *SOP transactions on Statistics and Analysis*, 1(3):1–15.

[61] Murthy K. (2004). Crop growth modeling and its applications in agricultural meteorology. *Satellite Remote Sensing and GIS Application in Agricultural Meteorology workshop 7-11 July. 2004*, pages 235–261.

[62] Natali H., Yatsenko Y., Goetz R.-U., and Xabadia A. (2009). A bang bang regime in optimal harvesting of size-structured populations. *Nonlinear Analysis*, 71:2331–2336.

[63] Ohrnberger P. (1999). Forestry research support programme for Asia and the Pacific, Bangkok, Thailand. Technical report, "The bamboos of the world": Annotated nomenclature and literature of the species and the higher and lower taxa IDRC, Ottawa, Canada.

[64] Ongugo P. O., Sigu G. O., et al. (2000). Production-to-consumption system: A case study of the Bamboo Sector in Kenya. Technical report, KEFRI, Kenya.

[65] Pasquali S. (2001). The stochastic logistic equation: Stationary solutions and their stability. *Rend. Sem. Mat. Univ. Padova*, 106:165–183.

[66] Paulo M. J., Stein A., and Tomé M. (2002). A spatial statistical analysis of cork oak competition in two Portuguese silvopastoral systems. *Can. J. For. Res.*, 32:1893–1903.

[67] Paydar Z., Huth N., and Snow V. (2005). Modelling irrigated Eucalyptus for salinity control on shallow water tables. *Australian Journal of Soil Research*, 43:587–597.

[68] Peper P., McPherson E. G., and Mori S. M. (2001). Equations for predicting diameter, height, Crown width and Leaf area of San Joaquin Valley Street trees. *J. Arboriculture*, 27(6):306–316.

[69] Qi L. and Mao X. (2006). Stochastic population dynamics under regime switching. *J. Math .Anal. Appl.*, 334:69–84.

[70] R Core Team (2017). A language and environment for statistical computing. R Foundation for Statistical Computing. `http://www.R-project.org/`.

[71] Rajeew K. and Chaturevdi S. (2009). Crop modeling: A tool for agricultural research. g.b. Pant University of Agr. & Tech. Pantnagar. `http://agropedia.iitk.ac.in/content/crop-model`. Accessed: 15-09-2017.

[72] Rao A. N., Dhanarajan G., and Sastry C. B. (1985). "Bamboo Research in Sir Lanka" in "Recent Research on Bamboos". Technical report, Proc. International Bamboo Workshop, The Chinese Academy of Forestry, IDRC, People's Republic of China.

[73] Ratkowsky D. A. (1983). *Nonlinear Regression Modeling*. Marcel Dekker, New York.

[74] Raychaudhuri S. (2008). Introduction to Monte Carlo Simulation. *Proceedings of the 2008 Winter Simulation Conference*, pages 91–100.

[75] Richards F. J. (1959). A flexible growth function for empirical use. *J. Exp. Bot.*, 10:290–300.

[76] Robert C. and Casella G. (2009). *Introducing Monte Carlo Methods with R*. Springer-Verlag.

[77] Robert C. and Casella G. (2013). *Monte Carlo statistical methods*. Springer Science & Business Media.

[78] Rogovchenko S. P. and Rogovchenko Y. V. (2009). Effect of periodic environmental fluctuations on the Pearl-Verhulst model. *Chaos, Solutions and Fractals*, 39:1169–1181.

[79] Rojo J. P. (1996). Bamboo resources of the Philippines. Technical report, 1ˢᵗ National Conference of bamboo, Sarabia Manor, Iloilo City, Philippine.

[80] Saniya V., Gohil A., et al. (2017). Crop yielding prediction applying regression: Forecasting wheat/rice yield for Anand district. *IJAERD*, pages 1–6.

[81] Schneider M. K., Law R., and Illian J. B. (2006). Quantification of neighborhood-dependent plant growth by Bayesian hierarchical modelling. *J. Ecol.*, 94:310–321.

[82] Seber G. A. F. and Wild C. J. (1989). *Nonlinear Regression.* Wiley, New York.

[83] Shah M. A. (2014). Stochastic Logistic Model for Fish Growth. *Open Journal of Statistics*, 4:11–18.

[84] Singh A., Goyal V., Mishra A. K., and Parihar S. S. (2013). Validation of crop system simulation model for direct seeded rice-wheat cropping system. *Current Science*, 104(10):1324–1331.

[85] Skiadas C. and Skiadas C. H. (2010). Development, Simulation, and Application of First-Exit-Time Densities to Life Table Data. *Communications in Statistics-Theory and Methods*, 39:444–451.

[86] Snow V., Houlbrooke D., and Huth N. (2007). Predicting soil water, tile drainage, and runoff in a mole-tile drained soil. *New Zealand Journal of Agricultural Research*, 50:13–24.

[87] Soñmez T. (2009). Diameter at breast height-crown diameter prediction models for *picea orientalis*. Technical report, Artvin oruh University, Faculty of Forestry, Artvin, Turkey.

[88] Soboleva T. K. and Pleasants A. B. (2003). Population growth as a nonlinear stochastic process. *Math. Comp. Mod.*, 38:1437 – 1442.

[89] Sönmez T., Ercanlı I., and Keleş S. (2009). A distance-independent basal area growth model for oriental spruce (picea orientalis (l.) link) growing in mixture with oriental beech (fagus orientalis lipsky) in the Artvin region, North-East Turkey. *African Journal of Agricultural Research*, 4(8):740–751.

[90] Suryawan H. P. (2018). Analytic Solution of a Stochastic Richards Equation driven by Brownian motion. *J. Phys.: Conf. Ser.*, 1097:012086.

[91] Taha A. H. (2007). *Operations Research, An Introduction.* Prentice Hall, India, 8 edition.

[92] Thornley J. H. M. (2000). *Plant and crop modeling: a mathematical approach to plant and crop physiology.* Blackburn Press, Caldwell, NJ.

[93] U.S. Department of Health and Human Services (2014). "the health consequences of smoking - 50 years of progress": A report of the surgeon general. atlanta:. Technical report, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

[94] Vanclay J. K. (1994). *Modelling forest growth and yield. Applications to Mixed Tropical Forests.* CAB International.

[95] Verhulst P. F. (1838). Notice sur la loi que la population poursuit dans son accroissement. *Corresp. Math. Phys.*, 10:113–121.

[96] Wallingford Z., A. B. W., and W. S. K. (2004a). Using crown condition variables as an indicator of forest health. *Can. J. For. Res.*, 34:1057–1070.

[97] Wallingford Z. S. J., Bechtold W. A., and Stolke K. W. (2004b). Using crown condition variables as an indicator of forest health. *Can. J. For. Res.*, 34:1057–1070.

[98] Weiner J., Kinsman S., and Williams S. (1998). Modeling the growth of individuals in plant populations: local density variation in a strand population of Xanthium Strumarium. *Am. J. Bot.*, 85:1638–1645.

[99] X. Z., Li W., and Wang K. (2013). Ergodic method on optimal harvesting for a stochastic Gompertz-type diffusion process. *Applied Mathematics Letters*, 26:170–174.

[100] Yoshioka H. (2019). A simplified stochastic optimization model for logistic dynamics with control-dependent carrying capacity. *Journal of Biological Dynamics*, 13(1):148–176.

[101] Zeide B. (1993). Analysis of growth equations. *For. Sci.*, 39:594–616.

[102] Zhang X., Shuai Z., and Wang K. (2003). Optimal impulsive harvesting policy for single population. *Nonlinear Anal.*, 4:639–651.

APPENDICES

.1 Some crop models reported in recent literature

Software	Details
SLAM II	Forage harvesting operation
SPICE	Whole plant water flow
REALSOY	Soya beans
MODVEX	Model development and validation system
IRRIGATE	Irrigation scheduling model
COTTAM	Cotton
APSIM	Modelling framework for a range of crops
GWM	General weed model in row crops
MPTGro	Acacia spp.and Leucaena Spp.
GOSSYM-COMAX	Cotton
CropSyst	Wheat & other crops
SIMCOM	Crop (CERES crop modules) & economics
LUPINMOD	Lupin
TUBERPRO	Potato & disease
SIMPOTATO	Potato
WOFOST	Wheat & maize, Water and nutrient
WAVE	Water and agrochemicals
SUCROS	Crop models
ORYZA1	Rice, water
SIMRIW	Rice, water
SIMCOY	Corn
CERES-Rice	Rice, water
GRAZPLAN	Pasture, water, lamb
EPIC	Erosion Productivity Impact Calculator
CERES	Series of crop simulation models
DSSAT	Framework of crop simulation models including modules of CERES, CROPGRO and CROPSIM
QCANE	Sugarcane, potential conditions
AUSCANE	Sugarcane, potential & water stress conds., erosion
CANEGRO	Sugarcane, potential & water stress conds
APSIM-Sugarcane	Sugarcane, potential growth, water and nitrogen stress
NTKenaf	Kenaf, potential growth, water stress

.2 Simulation sheet 1 (*bambusa*)

period t (months)	P(t) logistic -model	P(t) polynomial
1	1	1.399845
2	1	1.33122
3	1	1.329735
4	1	1.3923
5	1	1.515825
6	1	1.69722
7	1.000000001	1.933395
8	1.000000005	2.22126
9	1.00000002	2.557725
10	1.000000082	2.9397
11	1.000000334	3.364095
12	1.000001359	3.82782
13	1.000005529	4.327785
14	1.000022485	4.8609
15	1.000091442	5.424075
16	1.000371881	6.01422
17	1.001512353	6.628245
18	1.006149899	7.26306
19	1.024998995	7.915575
20	1.10144325	8.5827
21	1.40832181	9.261345
22	2.584261797	9.94842
23	6.308579745	10.640835
24	13.03446331	11.3355
25	17.88420651	12.029325
26	19.50520872	12.71922
27	19.89262733	13.402095
28	19.97709488	14.07486
29	19.99513182	14.734425
30	19.99896615	15.3777
31	19.99978048	16.001595
32	19.99995339	16.60302
33	19.9999901	17.178885
34	19.9999979	17.7261
35	19.99999955	18.241575
36	19.99999991	18.72222
37	19.99999998	19.164945
38	20	19.56666
39	20	19.924275
40	20	20.2347
41	20	20.494845
42	20	20.70162

.3 R-code for K-S test for GOF

```
> #R-code (KS test)
> require(graphics)
> data5=read.delim(file.choose(),header=T)
> head(data5)
  polinomial b_0.673 b_0.5 b_0.3 b_0.1 b_0.0673 b_0.03 b_0.01
1    1.10     1    1   1  1.12   1.66   5.07  9.89
2    1.03     1    1   1  1.15   1.77   5.31 10.02
3    0.95     1    1   1  1.19   1.89   5.56 10.15
4    0.97     1    1   1  1.24   2.04   5.82 10.28
5    2.02     1    1   1  1.30   2.20   6.09 10.42
6    1.93     1    1   1  1.37   2.39   6.37 10.55
> x <- polinomial
> y1 <- b_0.673
> y2 <- b_0.5
> y3 <- b_0.3
> y4 <- b_0.1
> y5 <- b_0.0673
> y6 <- b_0.03
> y7 <- b_0.01
> ks.test(x, y1)
    Two-sample Kolmogorov-Smirnov test
data: x and y1
D = 0.56667, p-value = 8.583e-09
alternative hypothesis: two-sided
> ks.test(x, y2)
    Two-sample Kolmogorov-Smirnov test
data: x and y2
D = 0.45, p-value = 1.058e-05
alternative hypothesis: two-sided

> ks.test(x, y3)
    Two-sample Kolmogorov-Smirnov test
data: x and y3
D = 0.4, p-value = 0.0001355
alternative hypothesis: two-sided
> ks.test(x, y4)
    Two-sample Kolmogorov-Smirnov test
data: x and y4
D = 0.15, p-value = 0.5095
alternative hypothesis: two-sided
> ks.test(x, y5)
    Two-sample Kolmogorov-Smirnov test
data: x and y5
D = 0.1, p-value = 0.9251
alternative hypothesis: two-sided
> ks.test(x, y6)
```

Two-sample Kolmogorov-Smirnov test
data: x and y6
D = 0.4, p-value = 0.0001355
alternative hypothesis: two-sided
Two-sample Kolmogorov-Smirnov test
data: x and y7
D = 0.566667, p-value = 8e-09
alternative hypothesis: two-sided

.4   R-code for fitting 5-parameter logistic model

```
#Plotting data

> data2=read.delim(file.choose(),header=T)
> attach(data2)
> head(data2)
  time culms
1    1  1.03
2    2  0.95
3    3  0.97
4    4  2.02
5    5  1.93
6    6  1.95
> plot(time,culms,main="Number of culms per clump",ylab="Number of
culms",xlab="Time (months)",cex.main=1.5)
>
> cdf=culms/25
> require(nplr)
> npara1=nplr(time,cdf)
Testing pars...
The 5-parameters model showed better performance
> npara1
Instance of class nplr

Call:
nplr(x = time, y = cdf)
weights method: residuals


5-P logistic model
Bottom asymptote: 0.05094609
Top asymptote: 0.9959336
Inflexion point at (x, y): 1.386704 0.6440885
Goodness of fit: 0.9966751
Weighted Goodness of fit: 0.9999784
Standard error: 0.02119463 0.01302192


> plot(npara1,cex.main=1.2,main="Growth in proportion of culms w.r.t
max=25")
> + ylab="Proportion of culms (c/25)", main="Custom plot",
cex.main=1.5)
> lines(x2, y2, lwd=5, col="seagreen4")

> plot(npara1, pcol="grey40", lcol="skyblue1", showEstim=.5, showInfl=TRUE,
+ main="Default 'nplr' plot", cex.main=1.5)

plot (npara1, showInfl = T, xlab="time (months)",cex.main=1.5,cex.lab=1.2, ylab="Prop. of
culms", main="Progression in number of culms")
```

.5 R-code for Chi-Square test for GOF

```
> data4=read.delim(file.choose(),header=T)
> attach(data4)
> head(data4)
  culmdist  Gammadist
1  0.0412 0.011015457
2  0.0380 0.008141570
3  0.0388 0.008802577
4  0.0808 0.125729544
5  0.0772 0.107261099
6  0.0780 0.111195694
> x=culmdist
> y=Gammadist
> chisq.test(x,y)

        Pearson's Chi-squared test

data:  x and y
X-squared = 3420, df = 3249, p-value = 0.01815
```