THE STATISTICAL ANALYSES OF A COMPLEX SURVEY

OF BANANA PESTS AND DISEASES

IN

UGANDA

JAPHETH N. NGOYA

(B.Sc. Hons)

Submitted in partial fulfilment of the
requirements for the degree of

MASTER OF SCIENCE

in the

School of Mathematics, Statistics, Computer Science and Information Technology
University of Natal

Pietermaritzburg, 1999

TABLE OF CONTENTS

CHAPTER 5. Illustration of Cultivar Differences using Correspondence Analysis

DECLARATION

The research described in this thesis was carried out in the International Institute of Tropical Agriculture (IITA- Regional Office in Kampala, Uganda) in conjunction with the Uganda National Banana Programme. The analysis of the dataset was carried out in the Department of Statistics and Biometry, University of Natal, Pietermaritzburg, South Africa from 1997 to 1998, under the supervision of Prof. G. P. Y. Clarke.

I declare that this study represents my own original work and has not been submitted in any form for any degree or diploma to any university. Where use has been made of the work of others it has been duly acknowledged in the text.

....

Japheth N. Ngoya

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to the International Institute of Tropical Agriculture, Plant Health Management Scientists (Under the leadership of Dr.C. Gold) and the Uganda National Banana Programme (Under the leadership of Dr. E. Karamura) for allowing model the data set for my masters thesis.

My main and university supervisor , Professor G.P.Y. Clarke, for his expert guidance and supervision. I am highly grateful to Prof. Clarke for many hours of assistance, discussion and editing of the manuscript. His constant encouragement and help has been a great source of inspiration and allowed me to complete the study within the proposed time limit.

My gratitude to Dr. S. Nokoe (Biometrician, IITA in Ibadan) for his encouragements and comments of the manuscript. My special thanks to Dr. C. Gold and Dr. P. Speijer for continued guidance on the biological interpretations of the analysis and particular their critical comments on the biological implications of the statistical models developed in this study.

I would like to pay my over whelming gratitude to my sponsors, The Rockefeller Foundation Office in Nairobi (In particular Dr. John Lynam) for providing the scholarship to undertake this study. Without them, this study may not have been possible. I am sincerely greateful.

I would like to thank my family for their encouragement and for all the years of support during my study period.

Lastly, for all those who participated in one way or the other in carrying out the survey.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

The importance of the banana in Uganda and rationale for the banana pests and disease survey.

1.0 Introduction.

Bananas are in the forefront as far as international trade in fresh fruits is concerned. In Uganda (Africa's leading producer & consumer) bananas are the most important staple food both for export and domestic consumption. The Highland cooking (*Musa* type AAA-EA) and the beer bananas (AAA-EA, ABB and AB) form the staple diet in East African highlands(C.S. Gold, Ogenga-Latigo, M.W., W.K. Tushemereirwe, I.N., Kashaija and C. Nakinga (1993)) and are the most widely grown in Uganda. Stover and Simmonds (1987), described East Africa as the secondary centre of banana diversity and highland cooking cultivars are unique to that region.

Plantains or starchy bananas comprise hybrids of *Musa balbisiana* and *Musa acuminata* as the two distinct wild species (Simmonds, 1966). Bananas (*Musa spp.*) comprise a genetically diverse crop consisting of diploid, triploids and tetraploids (Stover and Simmonds 1987). The group includes dessert, cooking, roasting and brewing bananas. The dessert bananas which form the basis of world trade are invariably triploids of the *Musa acuminata* genome group. The proportion of world banana production directed towards processing(dried fruit, puree and flour) is very small. Bananas do not process well due to their lack of acidity, and their aromatic constituents do not preserve well. However, despite these qualities, the availability of fresh fruit the whole year round makes processing and preservation unnecessary.

In Uganda, yield decline has been reported in traditional banana growing areas such as Mpigi, Luwero, Mukono and Iganga districts and has led to the replacement of cooking bananas with beer types and/or annual crops (e.g Cassava and Sweet potatoes) (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)) .

Banana production has been hampered by a number of constraints which include pest complex (weevils, nematodes and diseases) which causes yield decline and shortens plantation life (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993); INIBAP, 1986). Farmers have identified weevils and deteriorating soil fertility as their key problems but damage due to nematodes and pathogens were often attributed to the other factors (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993).

Banana pest infestations vary between neighboring farms and across regions (Gold *et al.* 1993; Sebasigari & Stover, 1988). Poor documentation into how ecological factors, cultivar selection and farm management factors affect pest status have led into several diagnostic surveys carried out by the International Institute of Tropical Agriculture (IITA) Plant Health Management Division and The Uganda National Banana Research Program (NBRP) to elucidate major production constraints in banana cropping systems. A rapid rural appraisal was conducted in 1991 to help focus survey objectives and to provide an understanding of farmer perceptions and management options (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993). The primary objectives of this survey were: 1. To map the distribution of banana pests in Uganda; 2. To determine pest status for key pests and diseases; 3. To explain the ecological factors and farm management practices which influence pest and disease distribution; and 4. to elucidate farmer perceptions which guide pest/disease management decisions.

1.1 Climatic requirement of the Banana

The banana is primarily a crop of the humid tropical lowlands which are areas characterized by less than 10° latitude; less than 100m altitude; not less than 19°C mean minimum temperature and more than 100 mm rain every month. There is some difference of opinion over physiological temperature thresholds for bananas but there is general agreement that 14°C is the minimum temperature for growth, 22°C the optimum temperature for the dry matter increase and flower initiation processes, 31°C the optimum temperature for leaf area increase, and that growth stops at 38°C (Robinson 1993).

In the humid tropics, mean temperatures are generally within the optimum growth range 22 to 31°C throughout the year. There is no evaporative stress on the plant, no chilling occurs, and irrigation is not necessary to improve production. In the cool subtropics or semi-arid tropics however, there is usually a period during the year when either cold winter temperatures or heat stress limits production potential. Rainfall is erratic and seasonal, and must be supplemented by irrigation for maximum production. However, there are many other factors determining banana growth suitability as we shall see later.

1.2 The Banana Plant; Importance in Uganda

Uganda is situated in the tropics and falls within the East African Great Lake region with other countries as Burundi, Kenya, Rwanda, Tanzania and Democratic Republic of Congo (formerly Zaire). Bananas are believed to have entered sub-Saharan Africa by multiple introductions between the first and sixteenth centuries A.D (Price, 1995; Karamura, 1997). In Uganda, there are more than 100 endemic highland banana clones (denoted AAA-EA to emphasize the group is unique to the region) belonging to triploids. The AAA group banana cultivars are regarded genetically as *Musa acuminata* (Simmonds 1966). A small group of cultivars belonging to genome groups AAB, ABB, AB regarded genetically as *Musa balbisiana* (Simmonds 1966) are scarce but increasingly grown in some farmers fields.

The highland cooking banana (*Musa* spp. type AAA-EA, Matoke group as it is commonly referred to in the region) is the most important staple crop in Uganda and some of the other countries within the East African great lakes region (e.g, Burundi, Rwanda) with Uganda being the region's leading producer and consumer of bananas. The crop is a key component in both food security and the agricultural sustainability of the region. Well managed banana stands often persist for 30 or more years, even under low input conditions. At the same time, an extended harvest period ensures food and income throughout the year (Gold C.S, Karamura E.B, Kiggundu, Bugamba and Abera (1998)).

Highland banana reduces soil erosion on steep slopes and provides principal sources of mulch for maintaining and improving soil fertility (INIBAP, 1986). Therefore the banana-based cropping systems have offered the most sustainable option under East African mid- and high elevation ecological conditions.

Most of the banana grower farmers in Uganda are in small scale category with farms of less than 0.5 ha, however, commercial farms exceeding 20 ha are also present in the region. Tothill, (1940) reported the first yield decline of highland banana which seems to have accelerated in the 1960s and 1970s. The yield decline prompted many farmers to adopt other crops in replacement of the highland banana or shift to exotic beer banana cultivars from the cooking ones. In addition commercial sources of banana have also shifted to non-traditional banana growing zones, driving up transportation costs.

The country-wide estimates of mean yields are 7 tons/ha and average yields of 4.6 tons/ha in central Uganda (Uganda Ministry of Agriculture, 1992) appear to suggest low productivity levels in many areas. The data also imply the existence of potential yield gaps between "well managed" stands and those receiving low inputs or less management attention, hence it is widely believed that productivity levels could be corrected through better management. Factors underlying the loss of banana sustainability in the central region have not been very clear (Gold C.S, Karamura E.B and Tushemereirwe, in press). Nevertheless, apprehension that some process might be afflicting growth in current production areas has raised concern about the future of the banana in Uganda. "Such fears were heightened by banana weevil (*Cosmopolites sordidus* Germar) and nematode (*Radopholus similis)* outbreaks in the mid-1980s causing yield losses of up to 100% in some regions (Masaka and Rakai districts)" (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga .,1993; Sengooba, 1986; Sebasigari and Stover, 1988). The Ugandan National Agriculture Research Organization gives the banana its highest priority for research attention.
It has been suggested that yield decline of highland cooking bananas in central Uganda has resulted from the combined effects of both social and biophysical factors. These include

population pressure, reduced labor availability, shift of farmer attention to other crops, removal of nutrients without replenishment, banana weevil problems and/or introduction of new pests(the nematode *Radopholus similis*, black *sigatoka*). The highly sustainable systems which have typified Uganda banana production depended upon the maintenance of adequate soil nutrient status through fallowing, carefully managed crop mixtures and recycling of nutrients. A combination of land pressure and off-farm flow of nutrients through sale of bananas contribute to declining soil fertility and a non-sustainable situation. This, in turn, may exacerbate pest status by weakening the plant and reducing its ability to resist or tolerate pest attack. Thus, there is concern about nutrient replenishment and long-term soil fertility in the commercial areas of Southwestern Uganda (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)). Up to 1991, there were no baseline data to document regional banana production levels, cultivar distribution, farming systems and production constraints. For example, the distribution and importance of key pests and diseases was unknown.

Against this backdrop, the Uganda National Banana Research Program and the International Institute of Tropical Agriculture have undertaken surveys to characterize banana-based farming systems in Uganda. Particular emphasis has been on assessing production constraints and elucidating the factors underlying geographical shifts of primary growing areas (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)).

1.3 The Survey: Site, farm sampling selection criteria.

A survey of 120 banana farms was carried out in 1991. Initially 24 sites (see figure 1.1) over the country were selected using a geographical information system (GIS) package developed by the IITA's agroclimatology unit and UNEP and CIAT demographic, topographic and climatic data bases (Japtap, 1993; C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)). Uganda banana growing regions were stratified on the basis of human population density, elevation and the length of rainy season . Local district agricultural officers were also involved in the identification of the banana growing areas. A grid (8km square) map was used to select twenty-one sites while three supplementary sites, showing features of particular interest (e.g. high elevation), were also selected (table 1.1 and figure 1.1.)

Table 1.1. The 24 study selected sites.

Site	District	Village	Elevation	Rainfall	Population
1	Kabale	Bukindia	1760-1830	Supplementary	Supplementary
2	Bushenyi	Mitooma	1510-1670	*	**
3	Bushenyi	Ryeru	1340-1420	*	**
4	Mbarara	Rukiri	1430-1460	*	*
5	Mbarara	Bubare	1360-1410	*	*
6	Mbarara	Rugaga	1430-1470	*	*
7	Kaborole	Buhesi	1520-1560	Supplementary	Supplementary
8	Rakai	Kagamba	1190-1330	*	**
9	Masaka	Matete	1200-1270	*	**
10	Masaka	Ntusi	1260-1290	*	*
11	Mpigi	Kabulasoke	1160-1200	**	**
12	Mpigi	Buwama	1180-1260	**	**
13	Mubende	Kitenga	1200-1215	**	*
14	Mubende	Bulera	1250-1310	*	**
15	Mubende	Madudu	1210-1280	**	*
16	Kibale	Nkooko	1080-1180	***	*
17	Kibale	Matale	1180-1240	***	*
18	Kiboga	Bukamero	1160-1200	***	*
19	Luwero	Nyimbwa	1230-1280	**	**
20	Luwero	Butuntumul	1130-1180	**	**
21	Mukono	Kayunga	1050-1070	**	**
22	Iganga	Bulongo	1070-1120	**	**
23	Mbale	Butiru	1250-1270	***	**
24	Kapchorwa	Kaseren	1820-1870	Supplementary	Supplementary

Rainy Months: * 3-5 mo. ** 6-8 mo . *** > 8 mo

Population Density: * < 50 / km^2 **> 50 / km^2

(Survey Methodologies for banana weevil and nematode damage assessment in Uganda (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993))

Five farms were randomly selected from those within a village containing 100 or more banana mats which were at least two years old. This was set to ensure the presence of a "banana system" and to provide adequate sample sizes. Pests are sampled at key phenological events (e.g flowering or harvest) and therefore some small holdings (farms) were unable to satisfy minimum samples sizes at any given visit. Study plots contained 100 to 250 mats and the entire farm selected served as a study plot. A random sample of Twenty five mats containing at least 3 different aged plants (preferably flowered, preflowered and peeper) were selected within the plots.

Plants within the same mat were numbered (figure 1.2) with the mother plant bearing the 1st number and the youngest plant bearing the highest number for purposes of monitoring the same plant infestation over time and possibly assessing different types of infestation over time on the same plant. The growth stage of each plant(Pre-flowered, flowered, and harvesting) was also observed and recorded. Plants on selected mats were monitored for plant growth, pest incidence, and yield .

Pest incidence assessment focused on damage due to weevils, nematodes and leaf spot diseases. For weevil infestation assessment focused more on the plants which had reached harvesting stage or could have toppled due to pests and disease pressure. In this case, a destructive random sampling was used. Nematode infestation assessment was carried out on those plants which had attained harvesting or flowered stage. The assessment of infestation due to leaf spot was carried out on plants which had not been harvested or toppled (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)). The survey was carried out in six visits to each farm; January, March, May, July, September and November. Where possible the same plant was assessed for nematode or leaf spot with an exception of weevil infestation in which case a plant had to be cut and corm damage assessed and this could only be done once.

Figure 1.1 Sampling sites in banana growing areas in Uganda

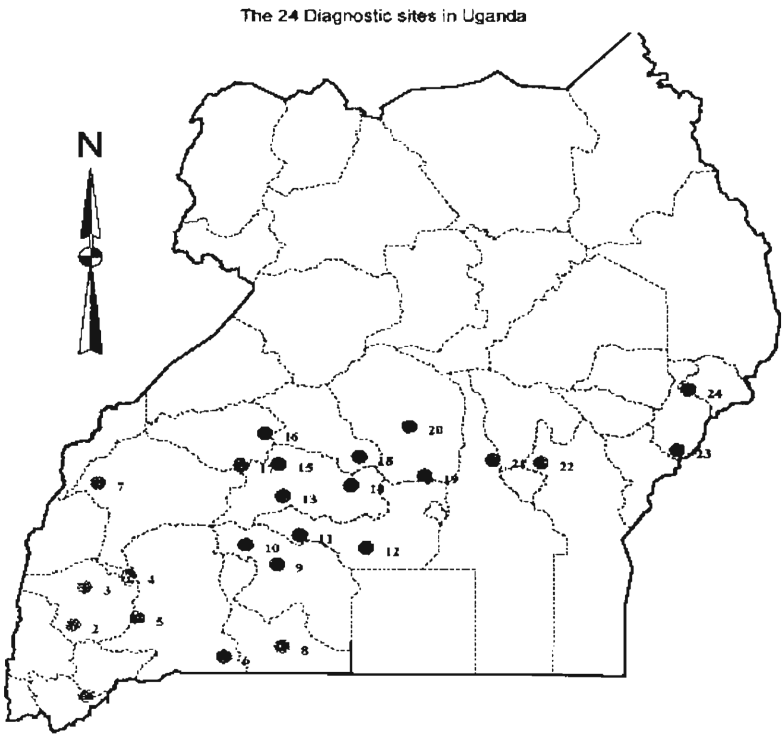
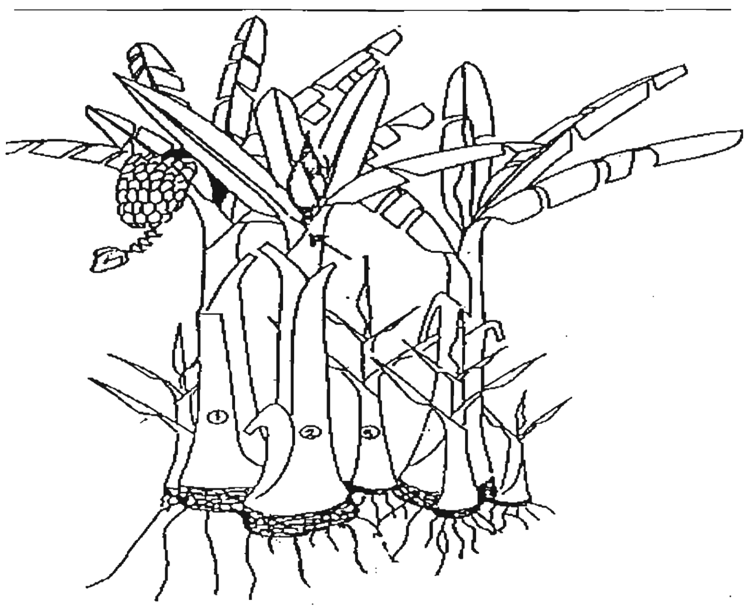


Figure 1.2 Banana Mat



1.4 The survey variables - Environmental and farm management practices

1.4.1 Environment factors.

Rainfall records for each month in each site averaged over a period of 10 years prior to 1991 were obtained from CIAT (Spanish name for "International Centre for Tropical Agriculture") research station based in Kampala, the rainfall figures were interpolated from existing data bases. Elevation(Altitude) measurements on each farm were carried out during the sampling visits using altimeter. Soil nutrients measurement on each farm were also obtained from the Makerere University. The data contains measurement of both 1st and 2nd depth but only the 1st depth measurement were used in this analysis since most of the soil nutrients composition is within the 1st depth. The soil variate(s) recorded were(:) pH, N, K, Ca, Na, Mg, Sand, Clay, Silt.

1.4.2 Farm Management practices

Farm management practices observed were intercropping, mulching, weeding, desuckering and de-leafing at farm level. Even though this data was recorded each time of a sampling visit, due to lots of inconsistence in each sampling visit, a final survey was undertaken after the pests and disease survey addressing the farm management practices and the data was incoporated with the weevil, nematode and leaf spot data sets. Identification of the sampled plants grouped into cultivars (varieties), genome group and genome sub-group was also carried out.

1.5. Survey Variables - Pests and diseases

Three major variables in these categories were examined namely damage by the banana weevil, by the root nematodes and by the leaf spot diseases as reported by (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)).

1.5.1 The Banana weevil.

Banana weevil, *Cosmopolites sordidus* (Coleoptera: Curculionidae) is an important pest of banana (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)). Weevils are distinguished by a pronounced snout and clubbed antennae arising midway on the snout (Figure 1.3). Weevils are herbivores and many species are pests. The banana weevil (*Cosmopolites sordidus)* is endemic to south/southeast Asia corresponding to center of origin of bananas but has been spread to all of the world's principal banana growing regions (Asia, Africa, Latin America, Pacific, Australia). This insect is specific to bananas and most widespread. It is economically the most important banana weevil, although its pest status in some areas is controversial. The weevil has 4 stages: egg, larva, pupa and adult and it's only the larval stage which causes damage to the banana plant. Banana weevil larvae bore in the corm, stem and pseudostem and represent the insect's most damaging stage. Young larvae tend to feed in the upper rhizome while older larva move throughout the corm, resulting to damage in the inner crossection, outer crossection and on the periphery of the banana corm. The larvae may also move through the rhizome from the mother plant into young suckers. In both East and West Africa, farmers have identified the banana weevil as a major cause of yield decline. However, damage symptoms are often confused by farmers with that caused by nematodes or leaf spots(diseases), *Fusarium* and soil deficiencies (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993). Banana weevil larval galleries (tunnels) weaken the plant and provide entry points for secondary pests(e.g fungal pathogens) which accelerate destruction and decomposition of rhizome tissues.

Figure 1.3: The banana weevil borer, *C. sordidus*

1.5.1.1 Measuring the banana weevil damage.

The weevil is an internal pest and requires destructive sampling. Therefore, sampled plants were either at the harvesting stage, dead, toppled or snapped. To assess weevil damage, each plant sampled was cut from the surface and cross sections were made at the base of the pseudostem and 5 cm below the base. For each cross section, weevil damage was assessed for the central and outer sections of the corm using the standardized scoring system. Figure 1.4 gives a sketch of method of scoring inner crossection and outer crossection.

Total damage in percentage for inner crossection for both lower and upper section of the corm was recorded under the variable inner crossection(xi), likewise outer crossection damage (xo), lower section(xl) and upper section(xu) damage were recorded . Variable xt(overall damage) was derived by averaging either xi and xo or xl and xu.

A modified Percent Coefficient of Infestation (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)) was used by scoring absence/presence of banana weevils infection in 20 sections. Ten 18° sections, each, from 0 to 5 cm and from 5 to 10 cm below to the base of the pseudostem were scored and totaled. This is a modification of the more traditional 10 section grid developed by Mitchell (1978).

Figure 1.5 explains the scoring method of Percent Coefficient of Infestation using a standard 10 segment template, and scoring 0 or 1 for absence or presence of infestation in each segment for both upper and lower sides of the template.

The measure of peripheral damage(dp) was calculated as follows. The periphery of the banana corm was pared between the cross sections at the collar and 10 cm below the collar, and the damaged area was recorded as percentage of the total area pared (C.S Gold, Ogenga-LatigoM.W., W.K. Tushemereirwe,I.N., Kashaija and C. Nakinga (1993)).

Figure 1.4 Cross section damage scoring system of weevil infestation.

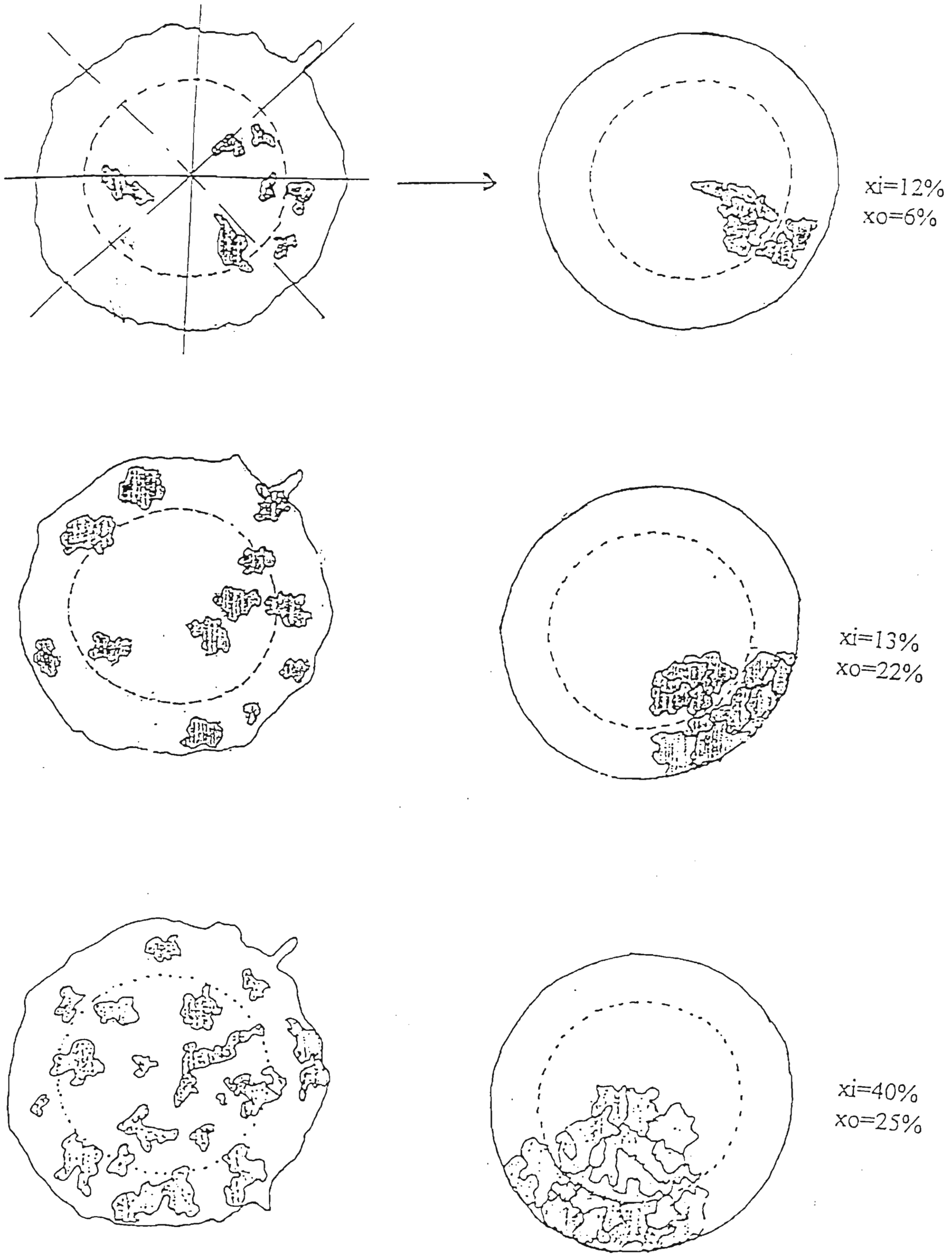
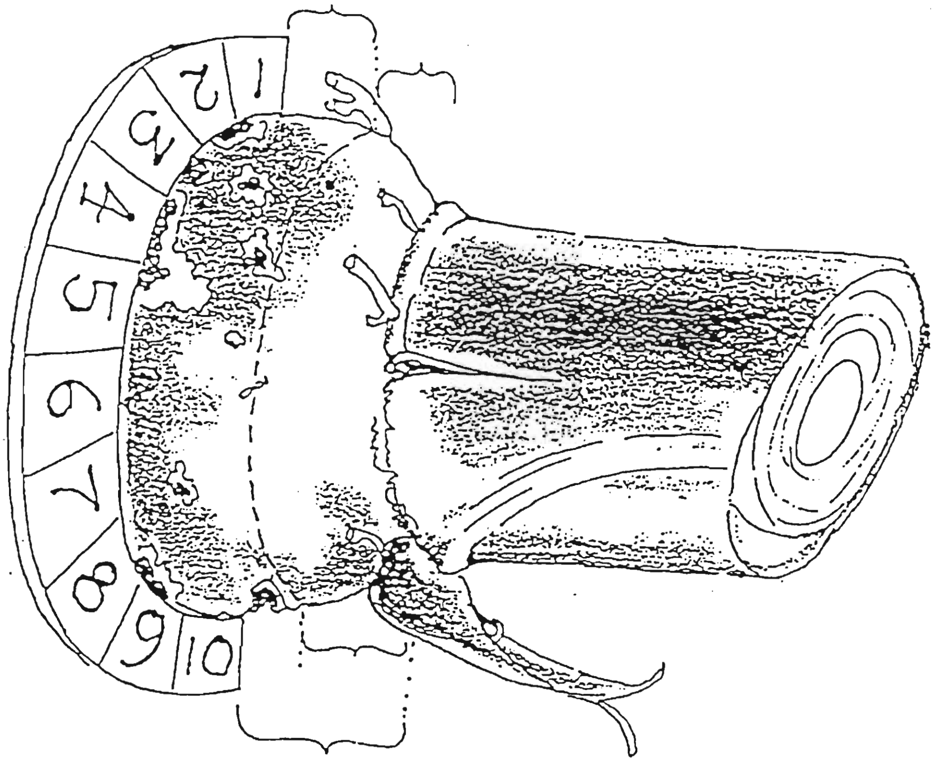


Figure 1.5 Percentage Coefficient of Infestation Scoring System (PCI).



	1	2	3	4	5	6	7	8	9	10	
U	0	1	0	1	0	0	0	0	0	0	2
L	1	1	1	1	1	1	1	1	1	1	10

1.5.2 The root knot nematode

Nematodes are microscopic organisms. As they are too small to be seen with the naked eye, the damage caused by these organisms is frequently underestimated. Banana and plantains are attacked by a complex of nematodes consisting of at least four general and numerous species(Speijer & de Waeli; 1997). Nematodes attack the root systems of a plant usually by penetrating at the tip of the root, although they are able to penetrate at any point. The earliest visible nematode damage on the outside of banana roots is elongated lesions that combine to form dark lesions over the entire root with raised cracks(Robinson J.C 1993). Inside the root, the nematode feeds by puncturing individual cells with its stylet and sucking out the cell contents.

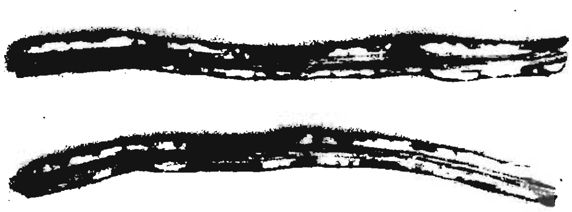


Figure 1.6. Corky red lesions caused by nematodes on banana roots.

As these cells are destroyed, the tissue changes in colour from normal healthy white to red, finally black (Figure 1.6)

The common nematode species in Uganda are *Radopholus similis, Pratylenchus goodeyi* and *Helicotylenchus multicinctus,* however plant toppling of the East African Highland cultivar (Musa AAA-EA) has been associated with high densities of *Radopholus similis* and *Pratylenchus goodeyi* (Speijer,P.R., Gold, C.S., Karamura, E.B., Kashaija, I.N.(1994b)). Nematodes move slowly within a plantation, but they are easily spread by water runoff, soil particles and planting material, to new plantings. The most important means of spread is via infested planting material and this is probably how the pest is introduced to most plantings (Robinson J.C. 1993). Damage to the roots by the nematodes can be further increased by other destructive organisms including fungi and bacteria e.g *Fusarium oxysporum f. sp. Cubense,* (Robinson J.C 1993) which enter through the lesions. The roots of a heavily-
infested plant subsequently rot away to short stubs that are unable to anchor the plant securely so that the latter topple over easily, especially those with bunches. Nematode damage on banana roots may be directly assessed in the field by determining the extent of necrosis on roots and rhizome for endoparasitic nematodes, or by determining the gall index on roots for root-knot. Indices for root health assessment include the percentage of functional roots, the extent of root necrosis, the extent of rhizome necrosis and nematode densities per 100 grams fresh root weight. In addition, observations on plant performance, including plant toppling, bunch weight, fruit filling and ratooning time, can provide insight into genotype susceptibility and nematode pathogen status(Stover 1972; Pinochet 1988).

1.5.2.1 Measuring root nematode damage.

Root necrosis reflects nematode damage resulting in premature root death. To measure the root necrosis (RI) , 5 roots were uprooted from the sampled plants which had attained either flowering or harvesting stage on each mat. The extent of damage was determined by employing a root necrosis (RI) index to the functional roots or by percent of dead roots from the sampled 5 sampled ones on each plant (Kashaija *et al.* 1994).

1.5.3 Leaf spot diseases.

Banana leaf spots in Uganda include black sigatoka(*M. Fijensis*), yellow sigatoka (*M. Musicola*) and cladosporium freckle (*Cladosporium Musae*). Black and yellow sigatoka produce similar symptoms and are hard to distinguish in the field although they can be readily separated from cladosporium freckle. The three diseases often occur on the same leaves although the interactions among them are not well understood(Tushemereirewe, 1997).

Yellow sigatoka begins as a yellow streak, up to 10mm long, parallel to the leaf veins, which darkens to form an elliptical brown spot. Over a few days the spots enlarge and the centre becomes grey (Figure 1.7). A yellow halo is common to older spots. The disease is favoured by humid and wet conditions and night temperatures over $21^{0}C$ (Robinson, 1993). It was first observed in Uganda in 1938 (Stover, 1972), had been considered a minor problem although its importance may have been underestimated (Tushemereirwe, 1996). Black sigatoka, first reported in Uganda in 1989 (Tushemereirwe and Waller, 1993), is considered one of the key constraints to banana production on a world wide basis. Yield losses can exceed 50% (Tushemereirwe and Waller, 1993). Black sigatoka causes a disease of leaves known as black-spot. The first sign is pin sized black spots near the leaf margins, usually on the old leaves under conditions of high humidity. These enlarge, assume an oval or spindle shape with a black edge and dead tissue in the middle. Midrib spots do not enlarge, but remains as black pin-spots. In warmer areas , the disease has killed young leaves although it is normally regarded as a parasite of older leaves. In Uganda the severity of this disease is estimated on recently flowered plants by identifying the youngest leaf with spots and the total area damaged(Tushemereirwe 1996).
The lesion can extend to the leaf margin forming a wedge of dead tissue (figure 1.7).

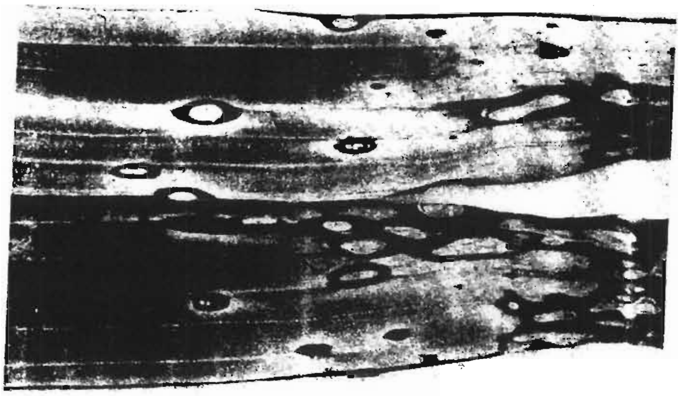


Figure 1.7 Black *sigatoka* lesions with a pronounced black edge and dead tissue in the middle.

1.5.3.1 Measuring Leaf spot infestation.

To assess leaf disease infestation, the leaf spot analysis was carried out on all the plants sampled and damage due to *sigatoka* (ST), *cladosporium (*CT), Youngest leaf with spots (YLS), youngest leaf with sigatoka(YLWS), youngest leaf with cladosporium(YLWC) and overall leaf spot damage (odt) were recorded. The variable YLS was identified as the most appropriate to use as indicator of leaf spot infestation due to its simplicity in identification and assessment and less likely to have subjective errors from one enumerator to the other as opposed to the other variables where sometimes is difficulty to differentiate leaf disease due to *Sigatoka* and *cladosporium.* It should be noted that the lower score (e.g score 1) the higher the leaf spot disease pressure while the lower the score (in our case 14) the less the leaf spot infestation. For those plants sampled and had no infestation, the score was a given a value of 15 instead of zero, for a score of zero will create a confusion on the averaging figures of site or farm infestation levels, by pooling the mean towards the average score of 1 which is actually misleading for this indicates high disease incidence. The score 15 for non infested plants was chosen after 14 which is the highest score within this variable and indicates very little leaf spot disease activity on the sampled plant.

CHAPTER 2

2.0 Introduction.

In this chapter we present overall the overview tables of distributions disease variates and associated of the sample sizes and general characteristics of variates in the survey. We particularly emphasize site distribution, cultivar distribution, soil characteristics averaged over site, rainfall and elevation and farm management practices across sites in relation to weevil, nematode and leaf spot infestation.

2.1 Environmental and farm management practices.

Overall the two most important environmental variables recorded are elevation(altitude) and rainfall. Table 2.1 shows the ranges of these two variates over the sample farms. It is important to notice the wide range of elevation from 1100 to 1800m. Table 2.1 illustrates the monthly rainfall distribution over the 24 sites with marked differences particularly between high elevation sites and others.

Tables 2.2 and 2.3 shows the soil characteristic and frequencies of farm management practices. Overall in the survey there were more than 60 different cultivars recorded on various farms. Some farms grew more than one cultivar. It was decided to concentrate on the 18 most common cultivars of the genome group AAA-EA (East African highland banana) which was in overall the most dominant genome group (over 85%) of the sampled plants. The selected 18 cultivars are those with frequency numbers over 100. The other cultivars were dropped from further analysis for the numbers were very low, some with total number of observations recorded in all six visits as few as 10 observations, and were scarcely distributed across sites.

Because the cultivar type is recorded for each individual plant in the survey, whenever a plant is sampled for either weevil, root necrosis or leaf spot damage, various plants were visited at various times(sometimes only once, sometimes twice ... six times) in the year. The fraction of plants recorded for each cultivar type will differ from records of weevils, root knot nematodes or leaf spot damage. Tables 2.4, 2.5 and 2.6 shows these figures. Overall the most commonly grown cultivar was Likhago with 9.1%.

Table 2.4, 2.5, and 2.6 also show how unbalanced the data set is with large differences in the popularity of various cultivars in different sites.

Tables 2.7, 2.8 and 2.9 show the number of plants sampled accross time and site. Evidently there is a very unequal distribution over time.

2.2 Plant pests and disease infestation.

Overall pests/disease infestation in each site is illustrated by table 2.10. Weevil infestation seems to be highly concentrated in sites of low and medium elevation and very little infestation on sites with high altitude. Nematode infestation (RI) , seems to occur on farms both with high and low altitude.

Table 2.1: Elevation and mean monthly rainfall of the sampled sites.

Site	Altitude	Rainfall (mm)												Mean
		Jan	Feb	Mar	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	
1	1796.23	68.00	89.00	116.0	155.0	98.0	27.0	16.0	51.0	97.0	110.0	126.0	95	87.3
2	1582.32	57.0	70.0	103.0	135.0	101.0	41.0	32.0	81.0	113.0	128.0	140.0	95.0	91.3
3	1380.24	39.0	54.0	92.0	132.0	98.0	47.0	31.0	76.0	106.0	118.0	124.0	73.0	82.5
4	1438.21	43.0	53.0	97.0	142.0	101.0	45.0	41.0	83.0	109.0	125.0	129.0	68.0	86.3
5	1387.55	45.0	57.0	94.0	120.0	82.0	26.0	22.0	58.0	91.0	108.0	119.0	71.0	74.4
6	1454.15	48.0	64.0	105.0	128.0	86.0	12.0	6.0	26.0	64.0	91.0	108.0	78.0	68.0
7	1536.34	46.0	61.0	121.0	169.0	121.0	66.0	56.0	109.0	143.0	164.0	144.0	80.0	106.7
8	1241.23	75.0	87.0	151.0	233.0	194.0	40.0	25.0	44.0	75.0	100.0	131.0	115.0	105.8
9	1241.23	42.0	49.0	90.0	122.0	87.0	26.0	22.0	51.0	87.0	99.0	103.0	65.0	70.2
10	1269.45	35.0	41.0	79.0	107.0	70.0	23.0	24.0	61.0	95.0	106.0	102.0	54.0	66.4
11	1179.58	44.0	50.0	94.0	131.0	91.0	36.0	32.0	61.0	92.0	113.0	112.0	68.0	77.0
12	1213.49	64.0	73.0	135.0	198.0	165.0	69.0	48.0	75.0	90.0	115.0	136.0	100.0	105.7
13	1205.89	35.0	43.0	87.0	130.0	87.0	38.0	44.0	87.0	112.0	131.0	115.0	57.0	80.5
14	1288.62	45.0	54.0	103.0	144.0	101.0	53.0	50.0	88.0	101.0	135.0	127.0	71.0	89.3
15	1245.78	35.0	43.0	87.0	130.0	87.0	38.0	44.0	87.0	112.0	131.0	115.0	57.0	80.5
16	1121.29	31.0	45.0	99.0	163.0	117.0	61.0	66.0	118.0	131.0	151.0	125.0	57.0	97.0
17	1207.87	36.0	446	103.	164.0	111.0	56.0	54.0	114.0	132.0	153.0	129.0	61.0	96.6
18	1183.76	31.0	48.0	95.0	159.0	123.0	61.0	73.00	117.0	122.0	146.0	119.0	57.0	95.9
19	1250.18	31.0	48.0	95.0	159.0	123.0	61.0	73.0	117.0	122.0	146.0	119.0	57.0	95.9
20	1160.38	41.0	57.0	107	159.0	123.0	70.0	66.0	106.0	112.0	136.0	122.0	70.0	97.4
21	1059.49	47.0	64.0	116	183.0	139.0	70.0	68.0	101.0	110.0	127.0	124.0	77.0	102.2
22	1098.27	43.0	62.0	116	199.0	155.0	76.0	77.0	114.0	114.0	127.0	119.0	72.0	106.2
23	1259.43	41.0	63.0	110	201.0	205.0	127.0	131.0	152.0	125.0	134.0	108.0	58.0	121.2
24	1847.64	30.0	50.0	89.0	168.0	182.0	117.0	153.0	165.0	111.0	119.0	91.0	44.0	109.9
Mean	1312.8	43.8	57.1	103	155.4	118.6	53.6	522	89.2	106.9	125.5	120.3	70.8	91.3
Q1	1190	35	48	93	130.5	89	37	28	61	93.5	111.5	113.5	57	57
Median	1255	42.5	54	101	157	106	50	46	87	110.5	127	120.5	69	91
Q3	1420	46.5	65.5	113	168.5	131	67.5	67	114	118	135.5	128	77.5	120.5

Table 2.2: Soil characteristics averaged over farm in each site

Site	Soil nutrients									
	Ph	N	K	NA	CA	MG	SAND	CLAY	SILT	OM
1	6.64	0.21	2.43	0.17	6.81	3.25	52.64	21.86	23.64	7.22
2	6.42	0.22	1.93	0.14	6.28	3.10	48.98	21.80	25.36	7.53
3	6.68	0.24	1.91	0.14	6.88	3.24	49.27	19.75	31.03	9.04
4	6.38	0.21	0.45	0.16	5.78	2.89	51.54	19.34	29.18	8.09
5	6.54	0.19	1.72	0.12	4.31	2.75	54.68	17.28	28.10	7.76
6	6.40	0.17	2.32	0.12	4.41	2.60	55.78	18.18	26.08	7.53
7	6.66	0.20	3.96	0.14	4.90	3.32	57.44	15.06	27.56	8.35
8	6.64	0.18	1.34	0.15	5.30	3.18	54.48	17.40	28.06	7.20
9	6.68	0.21	1.48	0.14	7.20	3.24	53.90	19.18	26.84	8.21
10	6.46	0.21	2.39	0.14	8.26	3.30	53.28	19.84	26.80	9.11
11	6.20	0.21	0.92	0.11	7.40	3.08	52.26	21.20	26.46	8.23
12	5.96	0.19	1.45	0.10	7.33	2.40	55.76	21.12	23.00	7.64
13	5.86	0.23	1.64	0.09	7.94	2.45	61.56	20.10	18.22	7.38
14	5.78	0.23	1.63	0.07	8.15	2.43	62.54	19.06	18.28	7.16
15	5.92	0.22	0.36	0.07	8.36	2.19	61.48	18.88	19.52	5.95
16	6.12	0.21	0.75	0.08	9.85	2.34	67.44	14.09	18.34	6.27
17	6.52	0.22	1.39	0.12	11.64	2.80	66.60	14.57	18.72	6.45
18	6.58	0.16	0.30	0.11	9.70	2.44	66.06	14.43	19.37	6.13
19	6.60	0.16	2.02	0.14	8.98	2.44	64.72	14.21	20.94	6.26
20	6.36	6.08	0.18	0.64	8.77	2.37	66.80	12.93	20.14	6.12
21	6.43	0.21	1.02	0.15	8.73	2.46	64.63	14.33	20.91	6.58
22	6.32	0.27	1.37	0.20	10.36	2.49	63.12	15.55	21.20	6.82
23	6.41	0.28	2.52	0.32	12.65	2.83	64.50	15.41	19.99	7.78
24	6.46	0.34	3.24	0.33	14.07	3.33	65.10	16.64	18.16	9.66
Mean	6.37	0.22	1.63	0.14	8.08	2.79	58.95	17.59	23.14	7.43
Q1	6.00	0.17	0.60	0.10	5.30	2.21	50.70	14	17.20	5.15
Median	6.4	0.20	1.15	0.12	7.75	2.60	62.40	18	21.50	7.95
Q3	6.8	0.27	2.40	0.17	10.50	3.28	66.40	21.3	27.50	9.27

OM - Organic Matter

<u>Units of the soil nutrients:</u>

N (Kg / ha^{-1}), K (Kg / ha^{-1}), Mg (Kg / ha^{-1}), Na (Kg / ha^{-1}), Ca (Kg / ha^{-1}), OM(%), Sand(%),Clay(%),Silt(%)

Table 2.3: Farm management practices in percentage form within sites.

Site	Intercroping		Mulching		Weeding		De-leafing		Desckuring	
	0*	1	0	1	0	1	0	1	0	1
1	40	60	0	100	0	100	0	100	20	80
2	20	80	0	100	0	100	20	80	20	80
3	23	77	0	100	0	100	19	81	0	100
4	40	60	0	100	0	100	20	80	0	100
5	40	60	0	100	0	100	20	80	0	100
6	20	80	0	100	0	100	20	80	20	80
7	20	80	0	100	0	100	0	100	40	60
8	20	80	0	100	0	100	0	100	40	60
9	20	80	20	80	20	100	0	100	60	40
10	20	80	20	80	20	80	0	100	80	20
11	20	80	20	80	20	80	0	100	60	40
12	20	80	20	80	20	80	0	100	40	60
13	40	60	20	80	20	80	0	100	40	60
14	40	60	0	100	0	100	0	100	20	80
15	60	40	0	100	0	100	0	100	0	100
16	80	20	20	80	0	100	0	100	0	100
17	100	0	20	80	0	100	0	100	0	100
18	78	22	21	79	21	79	21	79	21	78
19	60	40	20	80	40	60	40	60	40	60
20	40	60	40	60	40	60	40	60	40	60
21	20	80	41	59	38	62	37	63	40	60
22	0	100	62	38	38	62	38	62	38	62
23	0	100	83	17	17	83	17	83	37	63
24	20	80	80	20	0	100	0	100	0	100
Mean	35.04	64.95	20.29	79.70	12.25	88.58	12.17	87.83	27.33	72.62

* practice is not used

Table 2.4: Number of plants in each site over all six sampling visits categorised by variety - (selected for weevil infestation)

Site	Variety																			
	EN	KA	KB	KS	KI	LI	MB	MU	MV	NA	NT	NY	NW	NS	NI	ND	NF	SA	Others	Total
1	143	1					34		4			17				6			68	273
2	1		1			125	8				16	1			1	25			55	233
3		1				126	4		2		1				5	7			108	254
4						21			4	95					1	6			152	279
5			10			69	10	2	2	27	23				2	17			116	278
6			25			13	10	1	2	121					1	4			93	270
7		2		25						8		139				6			88	268
8			60		1	86	3	12	4	17	7					5			93	288
9			37			14	18	13	2	96	45				5	19		2	94	345
10		4	2		3	4	5	6	8	33	98				1	5			105	274
11			6	11	2	78	2	54	20		4	2		1	2	1	2	4	99	288
12		60	3	70		18	1	7	4		42			2			59		29	295
13		37		14							7	2		177			15	1	33	286
14		6	4	6	36			2		1	10	4	17			11	70	10	113	290
15		72			64			1	1		13			19			76		42	288
16		13	89		121			1	6	1			8	2	25	10	5		30	311
17				14			4		38	4	11	4	42	37				80	99	333
18		95					4	3		22		81				2		1	18	276
20		4	15			37	9	8	3	42	33		39		11	1	19	1	65	287
21		1			2		5	1		3	30		7		20	5	80		82	236
22		46					3			6	4	2	1		35	4	1		146	248
23						2	112	2		34	11				2				84	247
24						12	100	24											93	229
Total	144	342	261	188	229	598	247	146	108	551	452	254	124	238	121	136	333	102	1966	6540
%	2.2	5.2	4.0	2.9	3.5	9.1	3.8	2.2	1.7	8.4	6.9	3.9	1.9	3.6	1.9	2.1	5.1	1.6	30.1	100

<u>Codes used:</u>

EN - ENSENYI, KA-KAYINJA, KB-KIBUZI, KS-KISANSA, KI-KISUBI, LI-LIKHAGO, MB-MBWAZIRUME, MU-MUSAKALA, MV-MUVUBO, NA-NAKABULULU, NT-NAKITEMBE, NY-NAKYETENGU, NW-NAMWEZI, NS-NASSABA, NI-NDIIBWABALANGIRA, ND-NDIIZI, NF-NFUUKA, SA-SALALUGAZI, OTHERS

Table 2.5 Number of plants in each site over all six sampling visits categorised by variety - (Selected for root knot nematode infestation).

Site	Variety																			
	EN	KA	KB	KS	KI	LI	MB	MU	MV	NA	NT	NY	NW	NS	NI	ND	NF	SA	Others	Total
1	254	5				6	68		7	11		38			2	13			114	518
2			2			275	23		1	4	24	1	1			49			104	484
3		1				170	4	2		4	1				10	17			157	366
4	2	1			3	51			7	172					1	10	1		246	494
5			20			130	14	5	7	49	43				11	39			153	471
6			55		2	22	18	3	1	184			1		4	5			133	428
7		6		47					2	22					10	14			110	517
8		1	107		3	136	8	20	12	37	22			1		8			124	479
9			64		3	20	35	26	5	46	87				5	26		3	119	439
10		4	4		4	4	8	12	15	57	195	5			3	5			157	473
11			9	20	2	154	9	110	30	1	12	2		2	5	2	3	10	175	546
12		77		134	11	25	4	16	6		77		1	3	1		95	1	73	524
13		88		15		17				2	11	3		320		1	34	8	34	533
14		15	8	17	68			72		25	11	6	21		1	18	127	31	123	543
15		132		1	127	2		3	5	1	23			44			153		62	553
16		24	147		169		2	1	9	1			17	4	37	17	6	1	75	510
17				15			2	7	51	7	11	13	66	83		3		151	163	572
18		227					11	5		30	2	134				3		2	20	434
19		1	10	90		8	38	38	14	4	172	4	25		15	1	12	3	55	490
20		5	20			80	17	18	7	68	34		75		12	13	28	1	73	451
21		3		1	1		8	2		7	74		14		33	7	76		181	407
22		95					2	3		49	7	4	1		54	11			177	403
23							222	7		38	62			1					129	459
24								30		12	40								310	392
Total	256	685	446	340	393	1100	493	380	179	831	908	516	222	458	204	262	535	211	3067	11486
%	2.23	5.96	3.88	2.96	3.42	9.58	4.29	3.31	1.56	7.23	7.91	4.49	1.93	3.99	1.78	2.28	4.66	1.84	26 70	100

Codes used:

EN - ENSENYI, KA-KAYINJA, KB-KIBUZI, KS-KISANSA, KI-KISUBI, LI-LIKHAGO, MB-MBWAZIRUME, MU-MUSAKALA, MV-MUVUBO, NA-NAKABULULU, NT-NAKITEMBE, NY-NAKYETENGU, NW-NAMWEZI, NS-NASSABA, NI-NDIIBWABALANGIRA, ND-NDIIZI, NF-NFUUKA, SA-SALALUGAZI, OTHERS

Table 2.6 Number of plants in each site over all six sampling visits categorised by variety - (Leaf spot infestation).

Site	Variety																			
	EN	KA	KB	KS	KI	LI	MB	MU	MV	NA	NT	NY	NW	NS	NI	ND	NF	SA	Others	Total
1	237	7					83		12			27				12			112	490
2						313	14		1		43					28			108	507
3						267	1		2						12	16	5		209	512
4						47			11	191						15	4		233	501
5	1		12			114	13	5		46	33				10	33			170	437
6			51			29	14	6		212	2		1		5	10	1		195	526
7		10		27					1	21					7	4	1		138	517
8		1	104		4	131	8	15	14	40	16	308				3			170	506
9			69		1	35	44	31	3	78	82				2	28		1	144	518
10		5	7			21	3	9	13	48	230	5			6	7			166	520
11			12	19	3	151	4	82	36		11	2		1	10	5	84	4	183	523
12		104	3	132		26	4	17	5		72		1	3			33	1	73	525
13		94		24		7					32	1		283			138	8	30	512
14		15	2	11	61			70		10	2		12		2	21	154	43	139	526
15		149		1	84	1		3	10		24			33		23	2		65	547
16		28	177		154		1	1	9				22	5	41	38		1	80	559
17				16			6	3	59	5	5	14	70	61		3		162	162	566
18		265					13	1		30	1	151				1		1	24	487
19			22	64		17	27	34	16	1	128	4	34		19	2	26	22	102	498
20		7	30			66	18	11	8	96	52		66		13	9	31		69	476
21	1	4		2	1		8	2		14	72		14		30	4	29		296	477
22		101					4	5		14	1	15	1		80	12			282	515
23	1						229	7		16	90								141	484
24								42		28	58								336	464
Total	240	790	489	296	308	1225	494	344	200	850	954	527	221	386	237	274	508	223	3627	12193
%	1.97	6.48	4.01	2.43	2.53	10.05	4.05	2.82	1.64	6.97	7.82	4.32	1.81	3.17	1.94	1.94	4.17	1.83	29.75	100

Codes used:

EN - ENSENYI, KA-KAYINJA, KB-KIBUZI, KS-KISANSA, KI-KISUBI, LI-LIKHAGO, MB-MBWAZIRUME, MU-MUSAKALA, MV-MUVUBO, NA-NAKABULULU, NT-NAKITEMBE, NY-NAKYETENGU, NW-NAMWEZI, NS-NASSABA, NI-NDIIBWABALANGIRA, ND-NDIIZI, NF-NFUUKA, SA-SALALUGAZI, OTHERS

Table 2.7: Number plants sampled in each site categorised by time of sampling- (for weevil infestation)

Month of sampling

Site	Jan	March	May	July	Sept.	Nov.	Total (%)
1	32	42	50	50	50	49	273 (4.17)
2	9	26	49	48	51	50	233 (3.56)
3	2	50	47	51	53	51	254 (3.88)
4	30	46	49	50	54	50	279 (4.27)
5	37	42	49	50	50	50	278 (4.25)
6	30	38	50	50	52	50	270 (4.13)
7	22	44	50	51	50	51	268 (4.10)
8	41	47	50	50	50	50	288 (4.40)
9	35	48	50	51	51	50	285 (4.36)
10	32	42	50	50	50	50	274 (4.19)
11	42	50	50	46	50	50	288 (4.40)
12	42	48	50	50	53	52	295 (4.51)
13	50	50	37	50	48	51	286 (4.37)
14	42	48	47	50	53	50	290 (4.43)
15	35	51	50	50	52	50	288 (4.40)
16	50	45	49	51	66	50	311 (4.76)
17	50	53	48	51	79	52	333 (5.09)
18	18	27	47	50	39	45	226 (3.46)
19	48	37	42	50	47	50	274 (4.19)
20	50	41	46	50	50	50	287 (4.39)
21	26	39	30	50	44	47	236 (3.61)
22	8	45	45	50	54	46	248 (3.79)
23	27	35	36	47	52	50	247 (3.78)
24	20	40	39	47	42	41	229 (3.50)
Total	778	1034	1110	1193	1240	1185	6540
(%)	11.90	15.81	16.97	18.24	18.96	18.12	100.00

Table 2.8: Number plants sampled in each site categorised by month of sampling - (for nematode infestation)

Site	Jan	March	May	July	Sept.	Nov.	Total(%)
	Sampling months						
1	46	80	97	99	100	96	518 (4.51)
2	50	69	96	83	100	86	484 (4.21)
3	49	49	50	68	50	100	366 (3.19)
4	29	75	98	95	100	97	494 (4.30)
5	39	89	98	50	100	95	471 (4.10)
6	20	63	98	49	100	98	428 (3.73)
7	50	88	97	98	100	84	517 (4.50)
8	15	82	97	93	100	92	479 (4.17)
9	29	85	96	47	100	82	439 (3.82)
10	25	70	88	97	100	93	473 (4.12)
11	50	89	88	121	100	98	546 (4.75)
12	49	80	76	124	95	100	524 (4.56)
13	49	83	84	125	92	100	533 (4.64)
14	49	80	89	125	98	102	543 (4.73)
15	50	82	96	125	100	100	553 (4.81)
16	30	69	90	124	99	98	510 (4.44)
17	50	86	90	126	118	102	572 (4.89)
18	38	52	95	87	69	93	434 (3.78)
19	42	70	69	122	95	92	490 (4.27)
20	46	63	70	94	86	92	451 (3.93)
21	39	61	59	86	75	87	407 (3.54)
22	25	56	78	87	74	83	403 (3.51)
23	25	65	84	93	93	99	459 (4.00)
24	36	66	68	69	70	83	392 (3.41)
Total	930	1752	2051	2287	2214	2252	11486
%	8.10	15.25	17.86	19.91	19.28	19.61	100.00

Table 2.9: Number plants sampled in each site categorised by time of sampling- (for leafspot infestation).

Month of Sampling

Site	Jan	March	May	July	Sept.	Nov.	Total (%)
1	50	80	85	75	96	104	490 (4.02)
2	50	101	76	93	100	87	507 (4.16)
3	50	93	78	91	101	99	512 (4.20)
4	50	85	72	96	102	96	501 (4.11)
5	50	102	92	100	0	93	437 (3.58)
6	50	102	74	100	101	99	526 (4.31)
7	50	99	78	100	100	90	517 (4.24)
8	50	89	83	92	103	89	506 (4.15)
9	50	97	99	92	95	85	518 (4.25)
10	50	78	100	98	100	94	520 (4.26)
11	50	123	50	100	100	100	523 (4.29)
12	50	125	50	100	100	100	525 (4.31)
13	50	124	50	99	90	99	512 (4.20)
14	50	125	49	100	102	100	526 (4.31)
15	50	117	80	100	100	100	547 (4.49)
16	50	123	89	100	99	98	559 (4.58)
17	50	124	97	100	97	98	566 (4.64)
18	50	124	53	79	82	99	487 (3.99)
19	50	122	49	86	93	98	498 (4.08)
20	50	125	41	76	88	96	476 (3.90)
21	50	125	50	79	82	91	477 (3.91)
22	50	123	50	84	105	103	515 (4.22)
23	50	125	50	67	92	100	484 (3.97)
24	50	125	50	69	77	93	464 (3.81)
Total	1200	2656	1645	2176	2205	2311	12193
%	9.84	21.78	13.49	17.85	18.08	18.95	100.00

Table 2.10. Infestation due to weevils,nematodes & leafspot by sampling Sites

		Weevil infestation							Nematode	Leafspot
SITE	Altitude	xi	xo	xu	xl	xt	dp	pci	ri	yls
1	1796.23	0.00	0.02	0.01	0.01	0.01	0.10	0.12	11.34	14.40
2	1562.32	1.04	3.51	1.47	3.07	2.29	3.28	6.36	10.59	12.68
3	1380.24	1.13	3.88	2.08	2.93	2.53	4.31	9.63	7.15	8.75
4	1438.21	0.52	3.12	1.44	2.20	1.84	2.39	5.20	5.12	14.59
5	1387.55	1.09	4.67	1.88	3.88	2.90	2.29	5.36	6.20	12.43
6	1454.15	0.24	1.27	0.37	1.14	0.77	1.04	2.26	3.78	14.88
7	1536.34	1.35	3.70	2.33	2.72	2.55	3.52	7.14	7.89	14.30
8	1241.23	2.49	5.29	2.70	5.08	3.92	2.99	7.70	7.37	14.08
9	1241.23	1.83	3.21	2.41	2.63	2.54	2.37	5.46	3.98	14.48
10	1269.45	6.24	6.86	5.91	7.19	6.57	4.47	9.45	6.08	11.21
11	1179.58	3.90	5.43	3.90	5.43	4.69	4.25	9.83	3.40	7.33
12	1213.49	3.76	4.72	4.02	4.46	4.26	5.31	9.55	4.74	8.96
13	1205.89	2.78	5.74	3.37	5.14	4.28	5.65	1.21	4.71	8.88
14	1288.62	2.38	3.47	2.56	3.30	2.95	3.29	8.45	6.76	6.75
15	1245.78	1.07	2.09	1.41	1.75	1.60	2.38	5.66	9.82	10.08
16	1121.29	0.62	2.22	1.19	1.65	1.44	1.83	5.09	6.31	7.30
17	1207.87	5.89	6.49	5.17	7.21	6.22	6.83	3.63	9.59	5.15
18	1183.76	0.90	2.40	1.22	2.08	1.67	2.83	4.07	9.43	9.11
19	1250.18	1.87	4.66	2.58	3.94	3.29	3.93	7.72	5.37	5.20
20	1160.38	1.58	3.89	2.17	3.30	2.76	3.09	5.83	4.69	5.30
21	1059.49	2.99	4.97	3.46	4.50	4.00	4.89	8.61	5.71	5.10
22	1098.27	5.16	5.16	4.08	6.24	5.18	7.42	8.78	14.39	6.74
23	1259.43	0.44	2.53	1.42	1.55	1.51	2.74	5.07	2.45	10.99
24	1847.64	0.02	0.01	0.03	0.01	0.02	0.00	0.01	16.45	14.60
Mean	1312.80	2.10	3.77	2.43	3.45	2.96	3.41	6.88	7.15	10.09
Q1	1190.00	0.00	0.50	0.00	0.50	0.30	0.00	2.00	0.00	5.00
Median	1255.00	0.00	2.00	1.00	1.50	1.30	2.00	6.00	3.00	10.00
Q3	1420.00	2.00	4.50	2.50	0.00	3.50	4.00	11.00	10.00	15.00

CHAPTER 3

Basic principles behind the statistical techniques used in the survey.

3.0 Introduction.

Statistical analyses are aimed at:

(i) Quantifying various infestation factors on bananas due to pests and diseases.

(ii) Elucidating factors affecting pests (Weevils and Nematodes) / diseases (Leaf spots) distribution and their infestation levels.

(iii) Elucidating factors affecting the combined pests/disease infestation.

SAS/STAT (1989) or GENSTAT (1987) statistical programmes have been used in the analysis of the survey data, so as to look into the above objectives.

Because weevil infestation has been measured by several methods (xi, xo, xu, xl, xt, pci, dp), it would seem to be an ideal situation for multivariate techniques. Multivariate statistical methods focused on isolating most important weevil infestation variables, relationship of weevil infestation with environmental factors as well as farm management practices.

Due to the strongly hierarchical structure of the survey, mixed models are appropriate. Mixed model procedures examine the extent of weevil, nematode or leaf spot damage to banana plants with environmental factors and farm management practices as co-variates or factors. REML (Restricted Maximum Likelihood) analyses look at the extent of damage due to pests and diseases with site and farm effects as random effects, while environmental factors and farm management practices are fixed effects.

3.1 Multivariate methods.

Various methods were used on weevil infestation variates; Principal Component Analysis (PCA) was used to isolate the most important variables which could be used to determine the weevil infestation. Relationship of weevil infestation with environmental factors as well as farm management factors are examined by Cluster analysis and Multivariate analysis of variance (MANOVA). Logistic regression analysis was applied on the leaf spot data set with response variable YLS (youngest leaf spotted with disease) being our main focus as discussed in chapter 6.

3.1.1 Principal components Analysis (PCA)

Principal component analysis was originated by Pearson (1901) and later developed by Hotelling (1933). The application of principal components is discussed by Rao, C.R (1964), Cooley and Lohnes (1971), and Gnanadesikan (1977). Further statistical treatment of principal components is found in Krishirsagar (1972), Morrison (1976), and Mardia, Kent, and Bibby (1979). Principal components analysis finds orthogonal linear combinations of a set of variates that maximize the variation contained within them, thereby displaying most of the original variability in a smaller number of dimensions (Daultry, S. (1976); Rao, C.R. (1964)). It may be constructed using the sums of squares and products, or a correlation matrix, or a matrix of variances and covariances, formed from the data variates. This procedure was applied to the several variates which measure weevil infestation so as to identify the most important variate(s) displaying the maximum original variation in a smaller number of dimensions.

The principal components analysis produces matrices of columns of principal component loadings and scores corresponding to the latent roots. Each latent root corresponds to a single dimension, giving variability of the scores in that dimension. The loadings give the linear coefficients of the variables that are used to construct the scores in each dimension (Barlett, M.S. 1938)

In general, there are n observation and p variables. PCA then involves a rotation of the original axes.

Each component can be expressed as linear combination of the original axes:

$$U_j = v_{1j}X_1 + v_{2j}X_2 + \ldots + v_{pj}X_p$$

where j=1,2, ..., n ; such that:

(i) the first component has the maximum variation; $U_1 \geq U_2 \geq U_3 \ldots \geq U_p$

(ii) the second component is uncorrelated with the first and has maximum of remaining variation; $r(U_i,U_j)_{i \neq j} = 0$.

(iii) the third component is uncorrelated with the previous components and has maximum of remaining variation; and so on..

The equations provide the rules by which we can transform each unit $(x_1,x_2,\ldots x_p)$ to the new scales. U_1, U_2, etc are the principal component axes (latent vectors or eigen vectors), while v_{11}, v_{12} etc are the vector loadings or weights.

The loadings v's are usually scaled so that their sum of squares for each component equals 1.

In this way we transform to new uncorrelated components or variables which account for as much the variation as possible in descending order. It may be that the first two or three of these components (or new variables) account for "nearly" all the variation, say 85 or 90 per cent. We can say that the variation is represented approximately by the first two or three (principal) components and in favourable circumstances may be able to neglect the remainder (Morrison (1976)).

In this study PCA was applied to highly correlated (see chapter 4) weevil infestation variables so as to examines variables with principal components accounting most of variation. The variates are inner crossection (xi) , outer crossection (xo), upper crossection(xu), lower crossection(xl), peripheral damage (dp) all in percentage units and percent coefficient of infestation(pci) an index scaled from 0 to 20 which measures extent of outer damage on a banana corm.

3.1.2 Cluster analysis.

Classifying items into groups is a natural way of summarising the information in a large body of data. In some cases the items may possess a natural grouping. The aim of classification then would be to allocate a set of items to a set of naturally exclusive, exhaustive groups such that items within a group are similar to one another and items in different groups are dissimilar hence clustered or classified for description, data reduction or for topology/taxonomy(looking for some natural grouping). Fisher, R.A (1936) , showed that cluster methods can be categorised either as hierarchical or non-hierarchal methods merely assigning each item to a group. With hierarchical methods the groups themselves are arranged in a hierarchy; thus classification into, say, k groups if formed by splitting one of the groups in a classification of k-1 groups. The most widely used hierarchical methods act in an agglomerative rather than divisive way which are popularly displayed by use of the dendrogram. Non-hierarchical methods generally operate on a units by variate matrix and seek to partition the units into a specified number of groups to optimize some criterion (Gower, J.C (1967)). The most common clustering criterion used is maximizing the between groups sum of squares. Clustering is a good method of identifying optimum groupings (Anderberg, M.R (1973); Everitt, B.S. (1980); Coopper, M.C. and Millagn, G.W (1984)) and hence in this study was applied to group sites by their infestation levels.

3.1.3 Multivariate Analysis of Variance (MANOVA)

Multivariate analysis of variance, or MANOVA, for short, is treated extensively in most multivariate books and publications (Flury, B. (1951); Cole and Grizzle, (1966); Mead, R., Curnow, R.N and Hasted, A.M (1993); Huynh and Feldt, (1970);). Objectives pertaining to the explanation of social or physical phenomenon must be specified and then tested by gathering and analyzing data . In turn, an analysis of the data gathered by experimentation or observation will usually suggest a modified explanation of the phenomenon. Throughout the iterative learning process, variables are often added or deleted from the study (LaTour and Miniard, (1983)). Thus the complexities of most phenomena require an investigator to collect observations on many different variables. Richard A.J, Deab W.W (1982) in their book "*Applied Multivariate Statistical Analysis*" illustrates the use of MANOVA statistical methods to elicit information from these kind of data sets.

The model can be build into one equation:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon$$

where **Y** is *nxp*, **X** is *nxk*, β is *kxp*, and ε is *nxp*. Each of the *p* models can be estimated and tested separately as well as the *p* models can be tested simultaneously. We need to assume that the *p* dependent variables have errors that are independent across individuals but not across dependent variables for us to carry out multivariate tests.

In this study , MANOVA analysis is applied on highly correlated weevil infestations variables against several independent variables some being environmental factors as well as farm management practices. We however, point out that, due to hierarchical nature of the study(survey), this method does not take into account different error structures, e.g variance components due site, farm within sites random components. Therefore this method should be also taken as another measure of screening the factors affecting weevil infestation for further analysis as discussed in chapter 5 and 6.

Multivariate techniques have proven valuable in the fields of Medicine, Sociology, Business and Economics, Education, Biology, Environmental studies, Meteorology, Geology and Psychology (Hochberg, Y. (1974); Krishnaiah, P.R. and Armitage, J. V. (1966)). This shows multivariate methods are used in widely diverse fields.

Bernard Flury (1951) described multivariate analysis as "Mixed bag." It is difficult to establish a classification scheme for multivariate techniques that is both widely accepted and also indicates the appropriateness of the techniques. One classification distinguishes techniques designed to study interdependent relationships from those designed to study dependent relationships. Another classifies techniques according to the number of populations and the number of sets of variables being studied.

The classical MANOVA model allows us to partition the variation in the multivariate data matrix Y (n x p), which comprises observation on p variates from n individuals

3.2 Logistic Regression

Regression models for independent, discrete and continuous responses have been unified under the class of *generalised linear models*, or GLMs (Mc-cullagh and Nelder, 1989), thus providing a common body of statistical methodology for different type of response. Logistic regression is one of these models being used extensively for dichotomous response variables such as presence or absence of a disease (Agresti, A. 1984; Guthrie, D. 1981;Greenacre, M.J. (1984)). The logistic model assumes that the logarithm of the odds of a positive response is a linear function of the explanatory variables (Bishop, Y.M.M., Fienberg, S.E,, and Holland, P.W. (1985)), so that

$$\text{Log}\,(\Pr(Y_i=1)/\Pr(Y_i=0))=\text{Log}\,(\mu_i/(1-\mu_i)=X_i\,'\beta$$

Where Y is the bivariate response variable, X_i are the independent variables.. The regression coefficients, β , represents the change of the log odds of the response variable per unit change of X. The variance of the dichotomous responses is completely determined by its mean, $\mu_{i.}$ (Forthofer, R.N and Koch, G.G., (1973))

Specifically

$$\text{Var}(Y_i)=E(Y_i)\{1-E(Y_i)\}=\exp(X_i\,'\beta)/\{1+\exp(X_i\,'\beta)\}^2$$

This is to be contrasted with the linear model, where the $\text{Var}(Y_i)$ is usually assumed to be constant, σ^2, which is independent of the mean. In our study, logistic regression analysis is applied to the leaf spot diseases infestation in which the variable youngest leaf with spots (YLS) shows a 37.2% of the sampling units with no infestation against 62.8% with infestation (see table 6.17). Chapter 6 discusses in detail how the analysis is carried out.

3.3 The Mixed Model.

The standard linear model, or General Linear Model (GLM) is one of the most common statistical models:

$$y = \mathbf{X}\beta + \epsilon$$

In this expression, y represents univariate data, β is an unkown vector of fixed-effects parameters with known model incidence matrix $\mathbf{X}$, and ϵ is an unknown random error vector of independent random variables. The focus of the standard linear model is to model the mean of y by using the fixed effects β. The variance of each element of ϵ is assumed to be constant. The mixed model extends the general linear model by allowing a more flexible specification of the covariance matrix of ϵ. It allows for both correlation and heterogeneous variances, although still assuming normality (Searle, Casella, and McCulloch (1992).

Mixed models involve both fixed and random effects. Estimates of fixed effects are BLUE (best linear unbiased estimators) while prediction of random component means are BLUP(best linear unbiased predictor . Mixed models have a well known theoretical base(Henderson, 1984; Searle, S.R. 1971) and (Rubin. D.B, 1976; Self, S.G and Liang, K.Y. 1987). The mixed model effectively handles split-plot designs, repeated measures designs, and variances with heterogeneous structure. Winer (1971), Snedecor and Cochran 1980, and Milliken and Johnson 1992 gave an elaborate use of mixed model analysis on Split-Plot Designs.

The mixed model generalizes the standard linear model as follows:

$$Y = \mathbf{X}\,\beta + \boldsymbol{Zu} + \epsilon$$

Where X β represents the fixed part, ***Zu*** the random part with u being a vector random effects and ϵ is the unknown random error vector whose elements are no longer required to be independent. Estimates of the vectors β and $\boldsymbol{u}$ are to determined and;

Y is (n x 1) vector of observations

β is (p x 1) vector of fixed effects

$\boldsymbol{u}$ is (q x 1) vector of random effects with $\boldsymbol{u} \sim N(0, G)$

ϵ is (n x 1) vector of residual effects with the normality assumption N(0, R)

The generalization provided by the mixed model therefore enables not only the modeling of the mean of **Y** (as in the standard linear model), but modeling of the variance of Y as well.

To further develop this notion of variance modeling, assume that **U** and ϵ are uncorrelated and have expectations **0** and variances **G** and **R**, respectively. The variance of Y is thus

$V=\mathbf{ZGZ}' + \mathbf{R}$

Note that when $\mathbf{R}= \sigma^2\mathbf{I}$ and $\mathbf{Z}=0$, the mixed model reduces to the standard linear model.

The variances **G** and **R** may be estimated using the Method of Moments (MM), Maximum Likelihood(ML), the Minimum Variance (Norm) Quadratic Unbiased Estimated(MIVQUE) or the Restricted (Reduced) Maximum Likelihood (REML). In this study we only use REML.

By appropriately defining the model matrices X and Z, as well as the covariance structure matrices G and R, one can perform numerous mixed-model analyses using REML .

REML estimates the treatment effects and variance components in a linear mixed model: that is, a linear model with both fixed and random effects. REML is useful in situations where you would normally use ANOVA but have unbalanced data, or where you would normally use linear regression, but have more than one source of variation in the data.

REML is applicable in a wide variety of situations. It can be used to obtain information on sources and sizes of variability in data sets; analysing data from many fields, including identification of least reliable stages in an industrial process. REML also provides efficient estimates of treatments effects in unbalanced designs with more than one source of error. It can provide estimates of treatment effects that combine information from all the strata of a partially balanced design, or to combine information over similar experiments conducted at different times or in different places. You can thus obtain estimates that make use of the information from all experiments, as well as the separate estimates from each individual experiment (Robinson 1987). REML is suitable in the in analysis of data from complex surveys requiring small area estimation, longitudinal measurements from health and business research, spatial data, repeated measures, shrinkage estimates , split-plot designs and populations with unequal variances.

The method of residual maximum likelihood (REML) was introduced by Patterson and Thompson (1971) for a univariate analysis, this was further extended to multivariate analysis by Thompson (1973). It was developed in order to avoid the biased variance component estimates that are produced by ordinary maximum likelihood estimation: because maximum likelihood estimates of variance components take no account of the degrees of freedom used in estimating treatment effects, they have a downward bias which increases with the number of fixed effects in the model. This in turn leads to ignoring loss of degrees of freedom due to fitting the fixed effects and creates under-estimates of standard errors for fixed effects, which may lead to incorrect inferences being drawn from the data.

Thompson (1973) considered the case a multivariate two-way classification with treatments as fixed and blocks as random effects where the design and block structure were equal for all variates. The algorithm to estimate between-block and within-block(residual) variance and covariance components was presented very concisely using direct matrix products, vector, and general trace operators. The resulting equations were shown to be analogous to the univariate equations given by Patterson and Thompson (1971).

3.4 Correspondence Analysis

Correspondence analysis (CA) is a weighted principal component analysis of a contingency table. It finds a low-dimensional graphical representation of the association between rows and columns of a table. Each row and column is represented by a point in a Euclidean space determined from cell frequencies. Correspondence analysis is popularly used in France and Japan. In France, correspondence analysis was developed under the strong influence of Jean-Paul Benzécri; in Japan, under Chikio Hayashi. The technique apparently has many independent beginnings (for example, Richardson and Kuder 1933; Hirsfeld 1935; Horst 1935; Fisher 1940; Guttman 1941; Burt 1950; Hayashi 1950; Greenacre (1984). The algebra and geometry of correspondence analysis is provided in many statistical books (for example SAS / STAT (1990) , Volume 1) .

In this thesis, CA is used to reduce the dimensionality of tables in which rows are made up by cultivars and columns by measures of disease, adjusted by removing effects of rainfall and altitude (see Chapter 5).

CHAPTER 4

Multivariate Analyses

4.1 Principal Component Analysis(PCA) on Weevils infestation

The purpose of a PCA was to consolidate the numerous possible measures of weevil damage into one or two scores and relate these to other factors. Due to the hierarchical nature of the data, some analyses were carried out as observations averaged over all plants and farms in each site(called the site level), some were carried out averaged over plants within a farm (called the farm level).

4.1.1 Principal component analysis(PCA) at site level.

Standardization of all weevil variates were carried out first so as to equalize variance within the observed values. The transformed variables were then subjected to various analyses. Correlation analysis showed high significant correlation between the variates at visits. Table 4.1 shows correlation in the March sampling. Similar relationships were observed at all other sampling visits, implying suitability of principal component analysis. The results of PCA using site level data, are given in table 4.2, table 4.3 and figure 4.1. PC eigen value percentage contributions and coefficients of eigen vectors are very consistent over time and with almost equal indices for all variates in all the sampling visits. Over eighty percent contribution is attributed to the first eigen value (PRIN1) and most of the other percent contribution attributed to the second eigen value (Table 4.2). Looking at table 4.3, the first principal component for March sampling is $PRIN1=0.37xi+0.37xo+0.38xl+0.37xu+0.38xt+0.32pci+0.35dp$ where xi to dp are all standardized variables . The first PC is seen to be virtually a simple arithmetic average of all 7 variates and this accounts consistently over all sampling visits for over 82%. It must be noted that, PCA was not carried on the first sampling visit(January) due the fact that variable dp (peripheral damage) was only measured from the second sampling visit and onwards.

Table 4.1 Correlations for march sampling (for weevils infestation)

	xi	xo	xl	xu	xt	pci	dp
xi	1.00						
xo	0.82	1.00					
xl	0.96	0.95	1.00				
xu	0.92	0.92	0.91	1.00			
xt	0.94	0.96	0.98	0.97	1.00		
pci	0.70	0.71	0.72	0.72	0.73	1.00	
dp	0.42	0.40	0.35	0.48	0.40	0.45	1.00

Table 4.2 : Principal component on standardised variates at site level (for weevil infestation).

Components	March		May		July		September		November	
	Eigen-Value	%	Eigen-Value	%	Eigen-Value	%	Eigen-Value	%	Eigen-Value	%
PRIN1	6.40	83.0	7.01	87.7	7.24	90.5	7.09	88.6	6.90	86.3
PRIN2	0.77	9.6	0.49	6.2	0.38	4.8	0.41	5.2	0.47	5.9

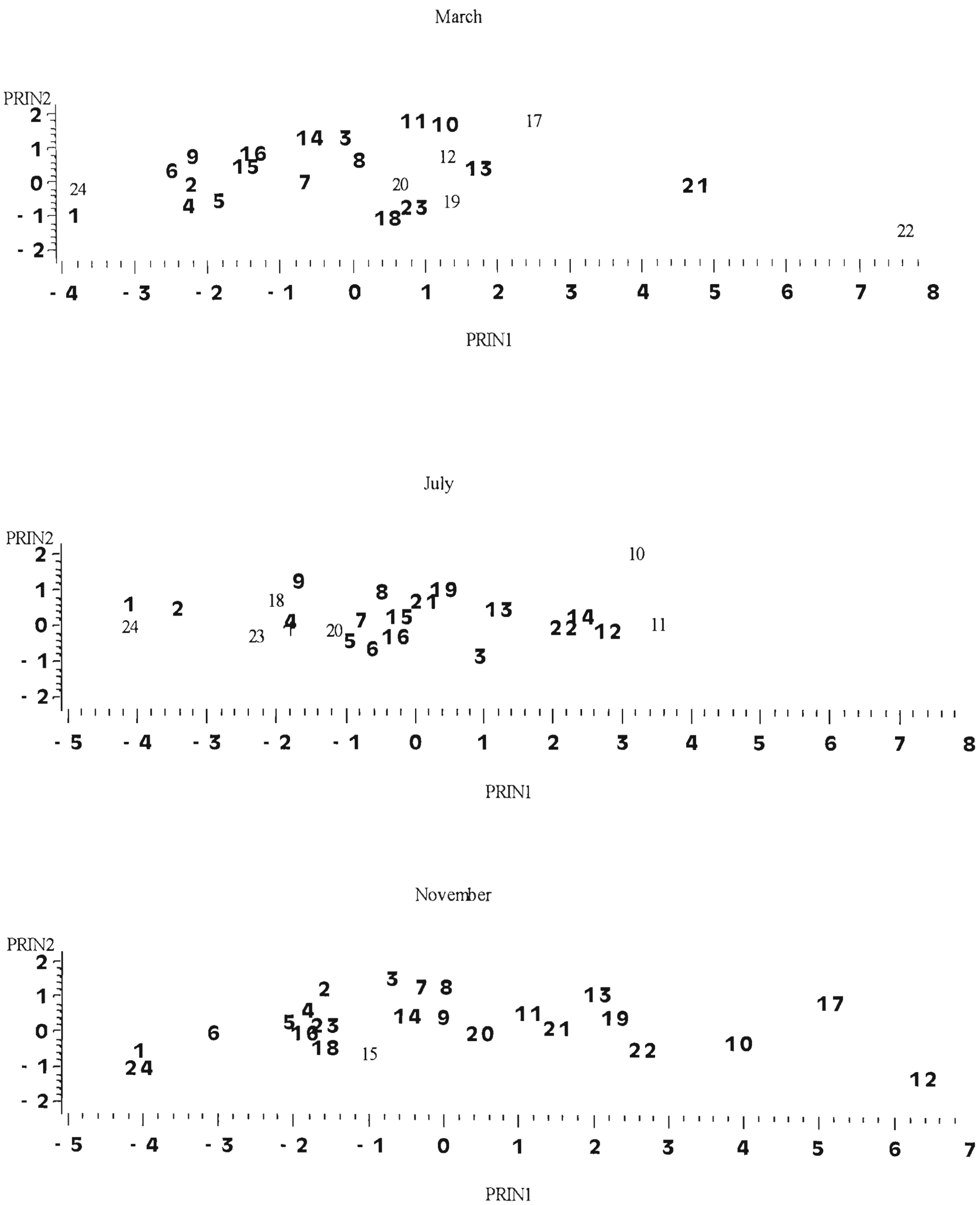
Table 4.3 : Coefficient of first two eigen vectors from principal components (for weevil infestation)

Infestation	March		May		July		September		November	
	PRIN1	PRIN2	PRIN1	PRIN2	PRIN1	PRIN2	PRIN1	PRIN2	PRIN1	PRIN2
xi	0.37	0.05	0.35	0.39	0.35	0.28	0.35	0.26	0.35	-0.50
xo	0.37	-0.22	0.36	-0.31	0.35	-0.27	0.34	-0.60	0.36	0.13
xl	0.38	-0.05	0.37	-0.04	0.36	-0.07	0.36	-0.24	0.37	-0.04
xu	0.37	-0.16	0.37	0.10	0.37	0.08	0.36	-0.12	0.36	-0.32
xt	0.38	-0.10	0.37	0.02	0.37	-0.01	0.37	-0.19	0.37	-0.17
pci	0.32	0.26	0.34	-0.28	0.35	-0.32	0.34	0.21	0.34	0.75
dp	0.35	-0.32	0.34	-0.48	0.35	-0.36	0.35	0.09	0.36	0.14

xi-Inner cross section, xo-Outer Cross section, xl - Lower cross section, xu-Upper cross section, xt -(xi+xo)/2, pci - Percent coefficient of infestation, dp- Peripheral damage,

A plot of the first and second principal components shows sites (1,6,24) high elevation appearing on the extreme left of each graph (figure 4.1) and low elevation sites (17,21,22) on the right hand side of each graph.

Figure 4.1. Plots of first principal component against second component on different cycles



Very similar trends in the grouping of sites by PCA are also given by the cluster analysis as shown in section 4.2.

4.1.2 Principal Component Analysis at farm level:

The results of PCA at farm level is shown in table 4.4, table 4.5 and figure 4.2 for weevil infestation variables. The first principal component accounts for over 75% of variance across all the sampling visits. The Eigen values and Coefficients of the variates are quite consistent over time just like at site level analysis. The infestation variables almost have equal weights as indicated in table 4.5. The indices on the variates seems to conform with the results at the site level analysis throughout all the sampling visits. The first principal component again appears to be a simple arithmetic average over all the variates as shown in table 4.5 over all the sampling visits. Figure 4.2 demonstrates the same pattern as figure 4.1 in site level groupings by altitude.

Table 4.4 First two principal components of weevil infestation at farm level for weevil infestation **(PRIN1 - First Principal Component, PRIN2 - Second principal component)**

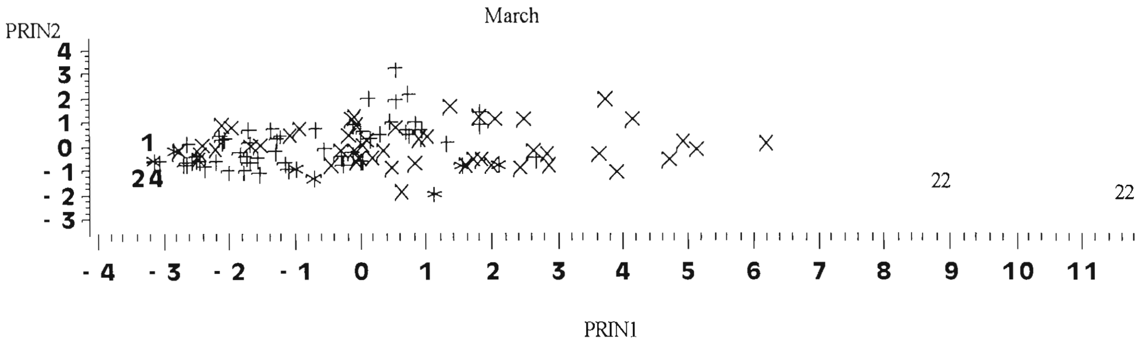
Components	March		May		July		September		November	
	Eigen-Value	%	Eigen-Value	%	Eigen-Value	%	Eigen-Value	%	Eigen-Value	%
PRIN1	6.15	76	6.5	81	6.7	84	6.5	82	6.4	80
PRIN2	0.78	9	0.6	7	0.5	6	0.6	7	0.5	7

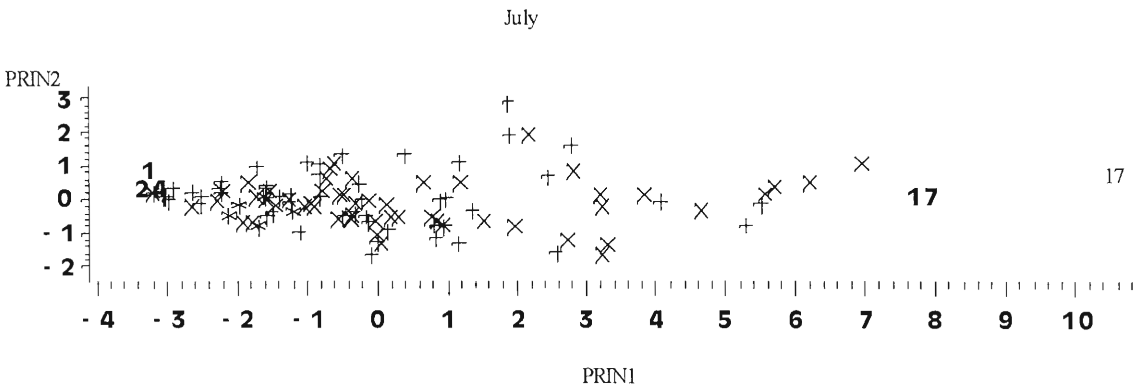
Table 4.5: Coefficients of the first two principal components(for weevil infestation)

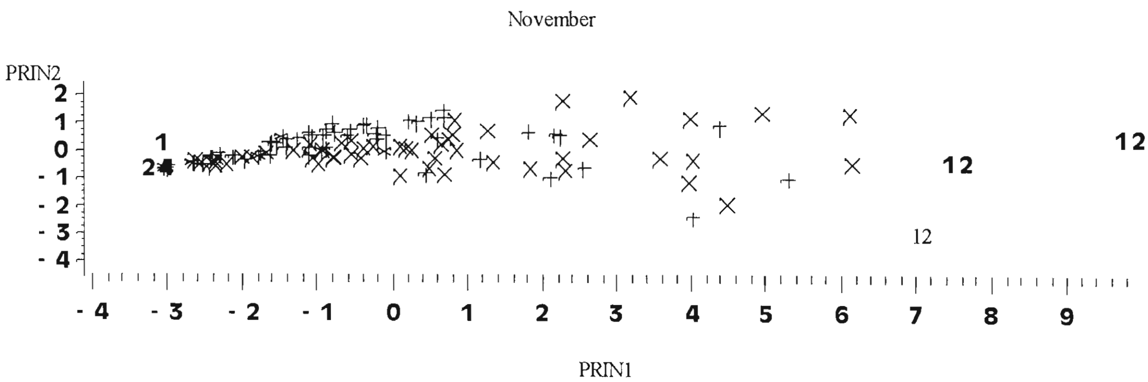
Infestation	March		May		July		September		November	
	PRIN1	PRIN2	PRIN1	PRIN2	PRIN1	PRIN2	PRIN1	PRIN2	PRIN1	PRIN2
XI	0.37	0.12	0.35	0.43	0.35	0.34	0.36	0.24	0.35	-0.44
XO	0.37	-0.28	0.36	-0.34	0.36	-0.27	0.35	-0.47	0.36	0.13
XL	0.37	-0.11	0.36	-0.03	0.36	0.01	0.37	-0.17	0.36	0.13
XU	0.36	-0.07	0.35	0.11	0.36	0.07	0.36	-0.09	0.34	-0.47
XT	0.39	-0.09	0.38	0.03	0.38	0.03	0.38	-0.14	0.38	-0.14
PCI	0.32	0.15	0.34	-0.32	0.34	-0.36	0.34	0.08	0.32	0.58
DP	0.35	-0.28	0.34	-0.44	0.34	-0.42	0.34	-0.08	0.35	-0.38

xi-Inner cross section, xo-Outer Cross section, xl - Lower cross section, xu-Upper cross section, xt -(xi+xo)/2, pci - Percent coefficient of infestation, dp - Peripheral damage

Figure 4.2 Plots of the first two principal components







4.2 Cluster Analysis

Table 4.6 shows results of cluster analysis minimising mahalanobis distances of six clusters. The clusters correspond well with the graph of PCs in figure 4.1. High elevation sites corresponds to grouping 1 and 2 and low elevation to grouping 5 & 6. It is quite clear that at sites of high elevation (over 1700m above the sea level) have less or few weevil attacks and those at low level elevation are prone to weevil damage as shown in Chapter 5.

Table 4.6 Cluster analysis at site level (for weevils infestation)

Cycle	Cluster	Groups	Sites
March	6	1	1,24
		2	2,5,15,16
		3	4,6,9
		4	3,7,8,10,11,14,20
		5	12,13,17,18,19,23
		6	21,22
July	6	1	1,6,24
		2	4,9,18,23
		3	2,5,7,8,15,16,19,20,21
		4	3,13,22
		5	10,11,12,14
		6	17
November	6	1	1,24
		2	6,16,18
		3	3,15
		4	4,5,8,9,19,20
		5	2,7,12,13,14,21,22
		6	10,11,17

4.3 MANOVA Analyses

An initial approach to the analyses of weevil damage was to use MANOVA. The idea here is to combine all 7 (xi, xo, xu, xl, xt, pci, dp) potential weevil infestation measures **(Y)** into one measure. MANOVA under SAS operates sequentially by fitting one explanatory factor at a time, using the optimal linear function of **Y** to separate effects of that factor. Then fit the next factor in the model until all listed factors have been fitted in order. At any stage a number of statistics are produced to measure significance. The statistic we use is the F ratio for Wilk's Lambda. Table 4.7 (a) and (b) shows these values. In this data set we introduce a new variable, **Prainfall**, which basically refers to the monthly rainfall figures for the month prior to the month of sampling. This is to investigate any relationship with this variable with weevil infestation and extended to nematodes and leaf spot infestation in chapters 5 and 6, for we believe rainfall prior to sampling month could have a significant effect on weevils although not found in the literature.

Three separate runs were made. Run 2 retained those variables from run 1 with $F>2$. Then run 3 retained those with F values > 3. In each run the order of variables was that from highest to lowest F value in the previous run. An examination of the coefficients used to calculate the best linear functions (table 4.8 (a) and (b)) indicates quite clearly that in this case, xi and xo contribute equally and none of the other measures make a serious contribution (xt, being ½(xi+xo)) make no additional contribution. Analyses were carried out on data transformed to the log scale and un transformed to see whether there might be some effects of heterogeneous variances.

Although there were 120 farms, we have here 1470 observation, implying that individual farm make several contributions (an average 12) to the data set. Also, this analysis ignores the hierarchical structure and hence MANOVA analysis does not account for various error structures e.g error structure due to site effect, farm within site effect, or even cultivars within the farm effects. The MANOVA and PCA analyses should only be seen as a means of preliminary screening. A more vigorous mixed model approach is used in chapters 5 and 6.

Table 4.7 (a) and (b)

(a) MANOVA analyses on weevil infestation variable before log transformation

Number of observations = 1470, No transformation on weevil infestation variables

Fitted terms	F-value for Wilk's lambda	Fitted terms	F-value for Wilk's Lambda	Fitted terms	F-value for Wilk's lambda
Elevation	14.32	Elevation	23.41	Elevation	27.84
Rainfall	4.68	Rainfall	6.38	Rainfall	7.37
Prainfall	1.53	Cultivars	4.42	Cultivars	4.43
Cultivars	4.41	Ca	3.39	K	3.98
Intercropping	1.27	De-leafing	1.72	Ca	3.75
Mulching	0.96	K	3.78		
Weeding	1.12	Na	1.76		
De-leafing	2.50				
Desuckering	1.52				
pH	1.33				
OM*	0.40				
N	1.68				
K	2.40				
Na	2.39				
Ca	3.09				
Sand	0.58				
Clay	1.35				
Silt	0.52				
Mg	0.95				

OM* - Organic Matter

(b) **MANOVA analyses on weevil infestation variable after log transformation**

Fitted terms	F-value for Wilk's lambda	Fitted terms	F-value for Wilk's Lambda	Fitted terms	F-value for Wilk's Lambda	Fitted term	F- value for Wilk's Lambda
Elevation	19.83	Elevation	24.84	Elevation	31.79	Elevation	32.97
Rainfall	5.54	De-leafing	3.99	Rainfall	8.74	Rainfall	7.58
Cultivars	5.17	Rainfall	8.05	Cultivars	5.24	Cultivar	5.30
Intercropping	1.32	Cultivars	5.26	K	5.26	K	5.40
Mulching	1.22	Na	2.24	De-leafing	3.93	De-leafing	6.04
Weeding	2.70	Ca	3.86	Ca	2.72	Weeding	3.2
De-leafing	6.27	K	4.84	Weeding	3.34		
Desuckering	1.55	N	2.49				
pH	1.53	Weeding	3.35				
OM*	0.30						
N	2.99						
K	3.68						
Na	3.98						
Ca	3.89						
Sand	0.80						
Clay	1.34						
Silt	0.66						
Mg	1.33						

Table 4.8 (a) and (b)

(a) **Characteristic vectors of $E^{-1}H$ from MANOVA for weevil infestation variables before log transformation..**

	xi	xo	xu	xl	xt	pci	dp
Elevation	-0.2924	-0.2956	-0.0014	0.0000	0.5900	0.0040	-0.00005
Rainfall	-0.1029	-0.1069	0.0065	0.0000	0.2046	0.0051	0.0006
Cultivars	-0.1173	-0.1159	-0.0006	0.0000	0.2358	0.0056	-0.0012
K	0.0278	0.0174	-0.0018	0.0000	-0.0374	0.0008	0.0016
Ca	0.2115	0.2052	0.0095	0.0000	-0.4215	0.0038	-0.0012

(b) **Characteristic vectors of $E^{-1}H$ from MANOVA for weevil infestation variables after log transformation**

	Log(xi+1)	Log(xo+1)	Log(xu+1)	Log(xl+1)	Log(xt+1)	Log(pci+1)	Log(dp+1)
Elevation	-0.0020	-0.0198	-0.0148	-0.0265	0.0707	0.0335	-0.0005
Rainfall	0.0334	0.0296	-0.0936	-0.1155	0.1530	0.0247	0.0037
Cultivars	0.0241	0.0452	-0.0099	-0.0081	-0.0263	0.0300	-0.0131
K	-0.0464	-0.1753	-0.0308	-0.0295	0.2839	0.0100	0.0109
De-leafing	0.0713	0.0537	0.0840	0.0394	-0.2273	-0.0011	-0.0022
Weeding	0.0565	0.0385	0.0761	0.0261	-0.1689	-0.0133	0.0109

Table 4.7 (a) and (b) shows elevation, rainfall and cultivars being a major factor in weevil infestation. The terms K, Ca, Na, N, de-leafing and weeding, produce inconsistent results, but sometimes significant contributions to these models. These lesser important factors (which also seem to be very strongly aliased with one another), are studied further in chapter 6.

The characteristic vectors shown in table 4.8 (a) and (b) seem consistent on both non transformed data and the transformed one, however, with some exceptions occurring. For example the rainfall effect on xi (-0.1029) for non transformed data and a positive effect of 0.0334, on the log scale.

CHAPTER 5

Illustration of Cultivar differences using Correspondence Analysis

5.1 The structure of the data

As explained in chapter 1, five farms from each of 24 sites were randomly selected for the survey. Each farm was visited on 6 occasions and disease incidence recorded on up to 10 plants on each farm. The ideal situation would have been to record the progress of disease on the same plant over time, but this was not possible. In practice, weevil infestation could only be measured when the plant was cut down at harvesting or when toppled. For this reason and other practical problems, it was very rare that the same plant was recorded for nematode, leaf spot and weevil infestation throughout the period. Consequently the data set ended up with various plants having been visited on various farms over time and no proper historical record at the plant level being available for a time series analysis.

A study of the survey results shows that altogether over 60 banana cultivars were planted on various farms, some with more than one cultivar falling into the survey from the same farm. It was decided to restrict attention to 18 of the most common cultivars (see table 5.1). Also, in order to follow a set of related plants over time, it was decided to average the records of the plants of the same cultivar on the same farm at the same sampling period.

Table 5.1 Frequency of selected cultivars before and after averaging*

CVNAME	*Before averaging* Frequency	Cumulative Frequency	*After averaging* Frequency	Cumulative Frequency
ENSENYI	144	144	23	23
KAYINJA	342	486	73	96
KIBUZI	261	747	106	202
KISANSA	188	935	76	278
KISUBI	229	1164	46	324
LIKHAGO	598	1762	190	514
MBWAZIRUME	247	2009	108	622
MUSAKALA	146	2155	85	707
MUVUBO	108	2263	74	781
NAKABULULU	551	2814	190	971
NAKITEMBE	452	3266	179	1150
NAKYETENGU	254	3520	61	1211
NAMWEZI	124	3644	53	1264
NASSABA	238	3882	61	1325
NDIIBWABALANGIRA	121	4003	78	1403
NDIIZI	136	4139	87	1490
NFUUKA	333	4472	111	1601
SALALUGAZI	102	4574	40	1641

* The averaging was carried out over plants of the same cultivar, sampled at the same time on the same farm

In this way the survey boiled down to 1641 sampling units, each unit now comprising the average record of plants from the same cultivar per farm per sampling period.

5.1.2 Defining the mixed model

Preliminary analysis indicates that major factors influencing disease are altitude, rainfall and cultivar. Two approaches have been used in this chapter to examine the effects of cultivar. Firstly a mixed model analysis with various error structures and with altitude, rainfall and cultivar as fixed effects has been carried out. Secondly the effects of altitude(elevation) and rainfall were removed and the residuals analysed through correspondence analysis to explain cultivar differences.

5.2 Measuring the weevil damage

5.2.1 Further data organization

The data set of 1641 sampling units was reduced further by removing data of the first sampling period (month 1) due to the fact that peripheral damage (dp) was not measured during this period. In addition to this, several outliers (8 observations) were identified (i.e measurements with weevil infestations of 30% or more) and deleted from the data set. This reduced the data set to 1470 sampling units.

5.2.2 Combining the various measures using Principal Component Analysis(PCA)

Due to skewness of the variates, all weevil measurement were converted to log(x+1) scale. Then a PCA using the correlation matrix was carried out (Table 5.2). The first principal component accounted for 82% of the total variation and the next only 9%. Scores using the first principal component were calculated for 1470 sampling units and these form the basis of further analysis.. It is worth noting that the first Principal component is very nearly a straight arithmetic mean of the 7 weevil variables (on a log scale) and consistent with Chapter 4.

Table 5.2 PCA analysis based on log transformed weevil variables.

Principal Component Analysis
1470 Observations
7 Variables

Simple Statistics

	XILOG	XOLOG	XULOG	XLLOG	XTLOG	PCILOG	DPLOG
Mean	1.0114	1.7331	1.3088	0.6698	1.1958	0.8480	1.0771
StD	0.7106	0.8634	0.7495	0.7331	0.7578	0.7132	0.7698

Correlation Matrix

	XILOG	XOLOG	XULOG	XLLOG	XTLOG	PCILOG	DPLOG
XILOG	1.0000	0.7092	0.8423	0.7975	0.8723	0.5919	0.6056
XOLOG	0.7092	1.0000	0.8472	0.9227	0.9586	0.7737	0.7722
XULOG	0.8423	0.8472	1.0000	0.7361	0.9067	0.6684	0.6715
XLLOG	0.7975	0.9227	0.7361	1.0000	0.9460	0.7285	0.7382
XTLOG	0.8723	0.9586	0.9067	0.9460	1.0000	0.7548	0.7626
PCILOG	0.5919	0.7737	0.6684	0.7285	0.7548	1.0000	0.8724
DPLOG	0.6056	0.7722	0.6715	0.7382	0.7626	0.8724	1.0000

Eigenvalues of the Correlation Matrix

	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	5.7243	5.0986	0.8178	0.8177
PRIN2	0.6257	0.3319	0.0893	0.9071
PRIN3	0.2937	0.0765	0.0419	0.9491
PRIN4	0.2172	0.0901	0.0310	0.9801
PRIN5	0.1270	0.1184	0.0181	0.9982
PRIN6	0.0086	0.0049	0.0012	0.9994
PRIN7	0.0036	.	0.0005	1.0000

Eigenvectors

	PRIN1	PRIN2	PRIN3	PRIN4
XILOG	0.3582	-.4864	0.4102	0.5530
XOLOG	0.3968	0.0112	-.4489	-.4065
XULOG	0.3754	-.3167	0.4130	-.6004
XLLOG	0.3891	-.0792	-.5691	0.3768
XTLOG	0.4108	-.1872	-.1651	-.0587
PCILOG	0.3545	0.5673	0.2302	0.0398
DPLOG	0.3567	0.5474	0.2351	0.1455

5.2.3 Analysis with cultivar as a fixed effect

The mixed model :

The standardized scores from the PCA were subjected to a mixed model analysis.

The Model used in the analysis was:

$$Y_{ijk} = \mu + \beta_{1i} X_{1i} + \beta_2 X_{2ij} + \tau_k + S_i + F_{ij} + C_{ijk} + M_{ijkl} \quad 3\text{-}1$$

where Y_{ijkl} represents the weevil score for month l averaged over all plants of cultivar k from farm j in site i.

The **fixed** terms in the model are:

X_{1i}, which denotes the average rainfall (over the past 10 years) during month l for site i

X_{ij} which denotes the altitude of farm j within site i.

and τ_k which denotes the effect of cultivar k.

The **random** terms in the model are:

S_i , which is the effect of site i, with variance σ_s^2 ,

F_{ij} , which is the effect of farm j in site i, with variance σ_f^2 ,

C_{ijk}, which is the effect of cultivar k from farm j in site i, with variance σ_c^2 ,

and M_{ijkl}, which is the effect of month l on cultivar k from farm j in site i, with variance σ_m^2 .

In this formulation, it is assumed that all random effects are independent. We refer to this as the "independence" model to distinguish it from the " correlated" model which has identical terms but for which successive month effects are correlated .

Table 5.3 shows the analysis with altitude(elevation), rainfall and cultivar as fixed effects. All three have substantial effects on weevil infestations. Among the cultivars Salalugazi and Nassaba appear to be more susceptible to weevil attack while Kisubi , Kayinja and to a lesser extent, Ndiizi show high resistance to weevil infestation.

5.2.4 Residual analysis

In order to examine cultivar differences more carefully, a model with cultivar as a random effect was fitted with only elevation and rainfall as fixed effects (Table 5.4) and the residuals where saved to be later analysed on cultivar effects with elevation and rainfall effects having been removed.

The residuals from this analysis were then examined using the SAS Univariate procedure. Table 5.5 gives the SAS output. Using the quartiles from table 5.5, all sampling units were divided into 4 classes was and cross tabulated by cultivar (see table 5.5). Correspondence analysis based on the frequencies of table 5.6 was carried to try to determine the cultivar differences with different levels of infestation.

Table 5.3 Reml analysis on weevils infestation with elevation, rainfall and cultivars as fixed effct

```
    9  vcomponent[fixed=elev+rain+cultivar;absorb=site] random=site,site/farm,site/farm/cultivar
   10  reml[print=model,effects,compo,deviance,wald;rmethod=all] wscore ;resid=res
***** REML Variance Components Analysis *****
              Response Variate : wscore (Weevil pc score)
              Fixed model      : Constant+elev+rain+cultivar
              Random model     : site+site.farm+site.farm.cultivar
              Number of units  : 1470
              Absorbing factor : site

              * Residual term has been added to model
              * All covariates centred
              *** Estimated Variance Components ***
              Random term              Component        S.e.

              site                        0.635        0.236
              site.farm                   0.437        0.099
              site.farm.cultivar          0.204        0.069
              *units*                     1.813        0.080

              *** Deviance: -2*Log-Likelihood ***   Deviance   d.f.
                                                     2720.00   1447
              *** Wald tests for fixed effects ***

                 Fixed term              Wald statistic      d.f.
                 elev                           20.5            1
                 rain                           22.6            1
                 cultivar                      257.4           17

              *** Table of effects      S.E

              Constant        3.118     0.5217
              Elevation      -0.004     0.0008
              Rainfall       -0.005     0.0011

              Table of predicted means for cultivars ****(with multiple range comparisons)
              Cultivar                          Means

              Salalugazi                      4.542 a
              Nassaba                         4.366 a
              Kibuzi                          3.854 a
              Namwezi                         3.776 a
              Kisansa                         3.744 a
              Likhago                         3.710 a
              Nakitembe                       3.703 a
              Nakabululu                      3.701 a
              Mbwazirume                      3.644 a
              Nfuuka                          3.608 a
              Musakala                        3.579 a
              Ndiibwabalangira                3.458 a
              Muvubo                          3.491 a
              Ensenyi                         3.118 ab
              Nakyetengu                      3.004 ab
              Ndiizi                          1.723 bc
              Kayinja                         1.531 c
              Kisubi                          0.870 c
              Standard error of differences :Average          0.3286
```

Table 5.4 Reml analysis on weevil damage scores to derive residuals

```
   9  vcomponent[fixed=elev+rain;absorb=site] random=site,site/farm,site/farm/cultivar
  10  reml[print=model,effects,compo,deviance,wald;rmethod=all]wscore ;resid=res
***** REML Variance Components Analysis *****
              Response Variate : wscore (Weevil pc score)

              Fixed model      : Constant+elev+rain
              Random model     : site+site.farm+site.farm.cultivar

              Number of units  : 1470
              Absorbing factor : site

              * Residual term has been added to model
              * All covariates centred

              *** Estimated Variance Components ***

              Random term                Component        S.e.

              site                           0.787       0.284
              site.farm                      0.358       0.111
              site.farm.cultivar             0.803       0.112
              *units*                        1.774       0.080

              *** Deviance: -2*Log-Likelihood *** Deviance   d.f.
                                                   2865.76  1464

              *** Wald tests for fixed effects ***
                 Fixed term                Wald statistic      d.f.

                 elev                           15.9             1
                 rain                           23.1             1

              * All Wald statistics are calculated ignoring terms fitted later in the model
              *** Table of      effects     s.e
                 Constant        3.318     0.1991
                 Elevation      -0.003     0.0009
                 Rainfall       -0.005     0.0011
```

Table 5.5 Distribution of residual scores for weevil damage

Univariate Procedure

Variable=WRES10 **(Weevil residual score)**

Moments

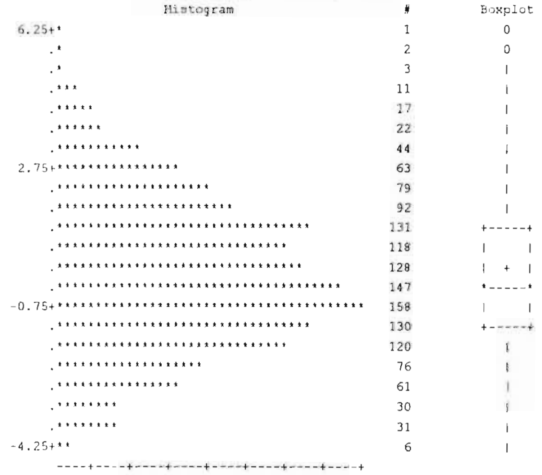
N	1470	Sum Wgts	1470
Mean	0.057207	Sum	84.094
Std Dev	1.899971	Variance	3.60989
Skewness	0.210636	Kurtosis	-0.34005
USS	5307.739	CSS	5302.928
CV	3321.233	Std Mean	0.049555
T:Mean=0	1.154408	Pr>|T|	0.2485
Num ^= 0	1469	Num > 0	710
M(Sign)	-24.5	Pr>=|M|	0.2104
Sgn Rank	6604	Pr>=|S|	0.6848

Quantiles(Def=5)

100% Max	6.078	99%	4.667
75% Q3	1.379	95%	3.26
50% Med	-0.08	90%	2.636
25% Q1	-1.342	10%	-2.3615
0% Min	-4.305	5%	-2.946
		1%	-3.859
Range	10.383		
Q3-Q1	2.721		
Mode	-3.821		

Extremes

Lowest	Obs	Highest	Obs
-4.305(	179)	5.214(	854)
-4.25(	475)	5.413(	265)
-4.25(	472)	5.777(	649)
-4.239(	768)	5.816(	1326)
-4.151(	1061)	6.078(	1383)



* may represent up to 4 counts

Normal Probability Plot

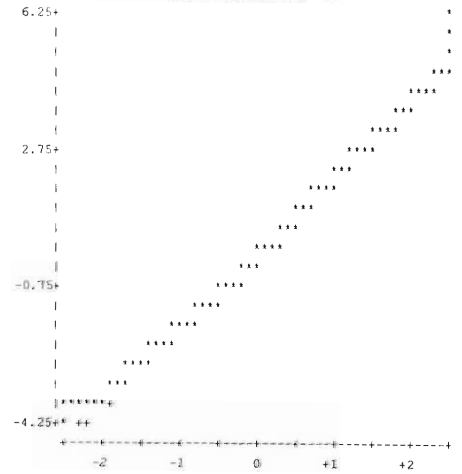


Table 5.6 Residual scores for weevil infestation by variety

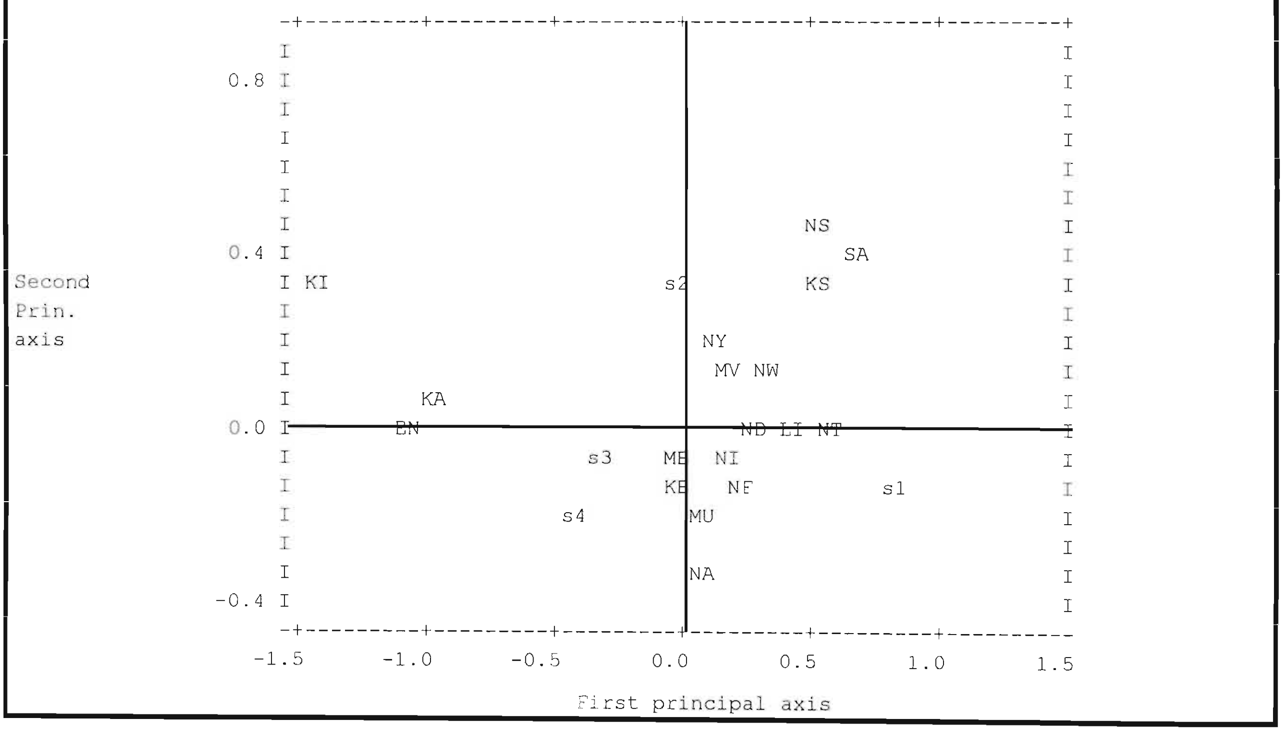
	Residual scores for weevil damage				
	<-1.3	>-1.3&<-0.1	>-0.1&<1.3	>1.3&<6.1	
Varieties	Frequency	Frequency	Frequency	Frequency	MEAN
KISUBI (KI)	36	3	1	0	-2.74
KAYINJA (KA)	42	14	3	4	-1.60
NDIIZI (ND)	55	18	7	3	-1.59
ENSENYI (EN)	15	5	0	1	-1.11
KIBUZI (KB)	25	27	23	19	-0.10
MBWAZIRUME (MB)	27	24	31	16	-0.08
NAKABULULU (NA)	33	67	41	32	-0.06
MUSAKALA (MU)	16	26	19	18	-0.03
NDIIBWABALANGIRA(NI)	13	21	13	23	0.27
NFUUKA (NF)	15	30	20	29	0.32
NAKYETENGU (NY)	13	8	18	17	0.33
MUVUBO (MV)	13	12	22	18	0.42
NAKITEMBE (NT)	24	43	46	51	0.43
LIKHAGO (LI)	21	44	58	49	0.46
NAMWEZI (NW)	7	9	14	14	0.55
KISANSA (KS)	6	8	25	27	1.13
NASSABA (NS)	6	4	15	28	1.30
SALALUGAZI (SA)	1	4	12	18	1.87

Table 5.7 Correspondence analysis of residual weevil scores by variety.

```
   5  corresp[print=roots,rowsc,colsc,rowin,colin]y \
   6          ;rowsc=rowsc;colsc=colsc
 *****  Correspondence analysis  *****
 ***  Squared singular values  ***
                  CA_Roots      %Roots     Cum%Root
            1      0.22808       81.84       81.84
            2      0.03906       14.01       95.85
            3      0.01156        4.15      100.00
 ***  Row Scores and Inertias  ***
                    rowsc
Pdata['rows']         1            2            3
            1      -1.4593      0.3606      0.0321
            2      -0.9945      0.0542     -0.0800
            3      -0.9870      0.0421      0.0114
            4      -1.1194      0.0292     -0.1426
            5      -0.0700     -0.1004      0.0287
            6      -0.0728     -0.0478      0.2057
            7       0.0387     -0.3254     -0.0121
            8       0.0568     -0.1807     -0.0285
            9       0.1346     -0.0562     -0.2172
           10       0.1796     -0.1161     -0.1613
           11       0.1217      0.2330      0.0986
           12       0.1650      0.1241      0.1357
           13       0.2477     -0.0129     -0.0318
           14       0.3035     -0.0292      0.0895
           15       0.2562      0.1041      0.0461
           16       0.4872      0.3042      0.0735
           17       0.4876      0.4871     -0.1919
           18       0.6583      0.3751     -0.1022
                  R_Inrtia
Pdata['rows']         1            2            3
            1     0.057947     0.003539     0.000028
            2     0.042387     0.000126     0.000275
            3     0.055004     0.000100     0.000007
            4     0.017900     0.000012     0.000290
            5     0.000313     0.000645     0.000053
            6     0.000354     0.000152     0.002822
            7     0.000176     0.012460     0.000017
            8     0.000173     0.001756     0.000044
            9     0.000863     0.000151     0.002247
           10     0.002064     0.000861     0.001664
           11     0.000564     0.002069     0.000371
           12     0.001204     0.000681     0.000815
           13     0.006844     0.000019     0.000113
           14     0.010777     0.000100     0.000937
           15     0.001965     0.000325     0.000064
           16     0.010658     0.004155     0.000243
           17     0.008573     0.008556     0.001328
           18     0.010319     0.003350     0.000249
  ***  Column Scores and Inertias  ***
                    colsc
Pdata['columns']      1            2            3
            1      -0.7859      0.1057      0.0033
            2       0.0270     -0.3301     -0.0495
            3       0.3334      0.0373      0.1691
            4       0.4268      0.1868     -0.1233
                  C_Inrtia
Pdata['columns']      1            2            3
            1      0.15461      0.00279      0.00000
            2      0.00018      0.02720      0.00061
            3      0.02782      0.00035      0.00716
            4      0.04547      0.00871      0.00379
```

Table 5.7 shows the results of the correspondence analysis. The first squared singular value explains 82% of the total variation, hence the results are largely explained by one axis alone. The row inertias for cultivars are dominated by cultivars Kisubi, Kayinja and Ndiizi, which are demonstrating high resistance to weevil infestation and occur in large numbers in the survey. From the column inertias it seems that the score category1 is dominant implying that for discriminatory purposes, scores 1 and 4 are important whilst 2 and 3 are far less so.

Figure 5.1 Correspondence analysis of weevil residual scores by variety



5.3 Measuring nematode damage

5.3.1 Data organization

Like the weevil data set described in the earlier part of this chapter, a data set of 1641 sampling units was created by averaging measurements taken from plants of the same variety in a farm in a sampling period and focusing on the 18 most common varieties. The nematode infestation measurement variable RI (root necrosis) was further transformed to log(RI+1) scale. The variable was then subjected to various statistical procedures.

5.3.2 Mixed Model analyses.

A mixed model analysis using REML in GENSTAT, was fitted with elevation, rainfall and previous month rainfall as fixed effects. The model fitted corresponded to model 3-1 with previous month rainfall added as an extra fixed effect. From this analyses it is quite evident that elevation, rainfall and previous month rainfall have a high influence on nematode infestation (see table 5.8).

However, as with the weevil scores, a model was fitted to remove effects of elevation, rainfall and previous month rainfall and residuals saved for further analysis. From this analysis cultivar differences are not statistically significant.

5.3.3 Residual analysis

Table 5.9 shows the mixed model output using REML in Genstat with elevation, rainfall and previous month rainfall as fixed effects, but cultivars as random effects.

The residuals were first subjected to the SAS univariate procedure so as to determine four classes of residuals scores using the quartile frequencies . Table 5.10 indicates a reasonable normal distribution with positive skewness.

Using the quartiles, a subdivision of 4 classes was established and cross classified by cultivars. Table 5.11 shows frequencies for the categories.

Frequencies from table 5.11 were then subjected to correspondence analysis. Table 5.12 and figure 5.2 gives the computer output.

The first and second squared singular values explain 85% of the total variation, thus the results are well explained by two axes. The lack of a clear trend in the column scores S1 to S4 confirms the small cultivar differences in nematode susceptibility.

Table 5.8 Reml analysis of fixed effects of elevation,rain, Prainfall & cultivar on nematode damage

```
vcomponent[fixed=elev+rain+prain+cultivar;absorb=site] random=site,site/farm,\
site/farm/cultivar
reml[print=model,effects,compo,means,deviance,wald;rmethod=all] rilog
***** REML Variance Components Analysis *****
Response Variate : rilog
Fixed model      : Constant+elev+rain+prain+cultivar
Random model     : site+site.farm+site.farm.cultivar
Number of units  : 1482 (159 units excluded due to zero weights or missing
values)
Absorbing factor : site
* Residual term has been added to model
* All covariates centred

*** Estimated Variance Components ***

Random term                Component        S.e.

site                          0.0980      0.0403
site.farm                     0.0811      0.0231
site.farm.cultivar            0.0149      0.0211
*units*                       0.7407      0.0318

*** Deviance: -2*Log-Likelihood ***   Deviance   d.f.

                                       1297.14   1458

*** Wald tests for fixed effects ***

   Fixed term              Wald statistic      d.f.

   elevation                        6.8           1
   rainfall                         4.6           1
   prainfall                       19.5           1
   cultivar                        12.4          17
*** Table of effects for ***
      Effects            Estimates      Stderr

      Constant             1.716       0.2801
      Elevation            0.0009      0.0003
      Rainfall             0.0008      0.0007
      Prainfall            0.0023      0.0005

*** Table of predicted means for cultivars ***
     Makitembe         1.775
     Namwezi           1.767
     Nfuuka            1.763
     Ensenyi           1.716
     Kisubi            1.711
     Kisansa           1.706
     Kibuzi            1.675
     Likhago           1.664
     Nassaba           1.655
     Mbwazirume        1.607
     Nakabululu        1.598
     Ndiibwabalangira  1.577
     Kayinja           1.568
     Musakala          1.566
     Nakyetengu        1.554
     Ndiizi            1.513
     Muvubo            1.476
     Salalugazi        1.464

Standard error of differences:     Average          0.1828
```

Table 5.9 Reml analysis on Nematode damage scores to derive residuals

```
vcomponent[fixed=elev+rain+prainfall;absorb=site] random=site,site/farm,\
site/farm/cultivar
reml[print=model,effects,compo,deviance,wald;rmethod=all]\
rilog ;resid=res
***** REML Variance Components Analysis *****
Response Variate : rilog

Fixed model       : Constant+elev+rain+prain
Random model      : site+site.farm+site.farm.cultivar

Number of units   : 1482 (159 units excluded due to zero weights or missing
values)
Absorbing factor : site

* Residual term has been added to model
* All covariates centred

*** Estimated Variance Components ***

Random term                 Component        S.e.

site                          0.0912      0.0373
site.farm                     0.0812      0.0225
site.farm.cultivar            0.0082      0.0199
*units*                       0.7442      0.0318
*** Deviance: -2*Log-Likelihood ***
                    Deviance   d.f.
                     1243.94  1475
Note: deviance omits constants which depend on fixed model fitted.

*** Wald tests for fixed effects ***

   Fixed term                Wald statistic      d.f.

   elev                             7.2            1
   Rainfall                         4.6            1
   prainfall                       19.4            1

* All Wald statistics are calculated ignoring terms fitted later in the model

*** Table of effects for Constant ***
     Effects          Estimates    Stderr

    Constant           1.639       0.0722
    Elevation          0.00098     0.00030
    Rainfall           0.00077     0.00070
    Prainfall          0.00242     0.0005
```

Table 5.10 Distribution of residual scores for nematode damage

```
Variable=RSCORE                      Univariate procedure
              Moments                                      Quantiles(Def=5)
N             1482  Sum Wgts       1482          100% Max    2.6118       99%     1.9207
Mean      -0.02068  Sum        -30.6438           75% Q3     0.7068       95%     1.4549
Std Dev   0.950313  Variance   0.903095           50% Med   0.03675       90%     1.1841
Skewness  -0.12346  Kurtosis   -0.74698           25% Q1    -0.7482       10%    -1.4612
USS       1338.118  CSS        1337.484            0% Min   -2.3049        5%    -1.6245
CV        -4595.92  Std Mean   0.024686                                    1%     -1.876
T:Mean=0  -0.83763  Pr>|T|       0.4024          Range       4.9167
Num ^= 0      1482  Num > 0         766          Q3-Q1        1.455
M(Sign)         25  Pr>=|M|      0.2031          Mode       -1.4539
Sgn Rank   -5300.5  Pr>=|S|      0.7478
                                 Extremes
                         Lowest    Obs     Highest    Obs
                        -2.3049(    457)   2.1882(     244)
                        -2.3049(    456)   2.2722(      56)
                        -2.0414(    521)   2.2833(     370)
                        -2.0237(    530)   2.3208(     925)
                        -2.0237(    528)   2.6118(    1435)
                              Missing Value         .
                              Count               159
                              % Count/Nobs       9.69
                         Histogram                         #             Boxplot
     2.7+*                                                  1               |
        .                                                                   |
        .*                                                  3               |
        .**                                                 6               |
        .***                                                9               |
     1.7+*********                                         25               |
        .**************                                    42               |
        .*******************                               55               |
        .***************************                       81               |
        .*********************************                 98               |
     0.7+****************************************         120            +-----+
        .******************************                    90            |     |
        .*****************************************        122            |     |
        .**************************************           114            *-----*
        .*********************************                 99            |  +  |
    -0.3+***************************************          117            |     |
        .************************                          72            |     |
        .*******************************                   91            +-----+
        .****************************                      83               |
        .*************                                     38               |
    -1.3+*******                                           20               |
        .****************************************         118               |
        .******************                                52               |
        .*******                                           19               |
        .**                                                 5               |
    -2.3+*                                                  2               |
         ----+----+----+----+----+----+----+----+-
         * may represent up to 3 counts
                                  Normal Probability Plot
          2.7+                                                   *
             |
             |                                                  +*
             |                                                ++ *
             |                                              ++***
          1.7+                                            +****
             |                                         +****
             |                                       ****
             |                                    ****
             |                                  ***
          0.7+                               ****
             |                              **+
             |                            ***
             |                          ***
             |                        ***
         -0.3+                      ***
             |                    ***
             |                  +**
             |                +**
             |              +***
         -1.3+            ++ *
             |          ******
             |     ******
             |******++
             |*  ++
         -2.3+*++
              +----+----+----+----+----+----+----+----+----+----+
                  -2        -1         0        +1        +2
```

Table 5.11 Residual scores for nematode infestation categorized by cultivar

Cultivars	Residual scores for nematode infestation				MEAN
	<-0.7	>-0.7&<0.03	>0.03&<0.7	>0.7&<2.6	
	Frequency	Frequency	Frequency	Frequency	
MBWAZIRUME(MB)	42.00	17.00	37.00	12.00	-0.23
MUSAKALA (MU)	31.00	15.00	34.00	5.00	-0.23
NDIIZI (ND)	32.00	12.00	36.00	7.00	-0.18
NAKABULULU (NA)	81.00	38.00	58.00	13.00	-0.14
MUVUBO (MV)	33.00	10.00	29.00	2.00	-0.13
KIBUZI (KB)	36.00	17.00	43.00	10.00	-0.07
LIKHAGO (LI)	52.00	35.00	92.00	11.00	-0.04
KISANSA (KS)	25.00	13.00	37.00	1.00	-0.03
NDIIBWABALANGIRA (NI)	28.00	5.00	41.00	4.00	0.02
ENSENYI (EN)	7.00	5.00	10.00	1.00	0.02
NAKITEMBE (NT)	55.00	32.00	85.00	7.00	0.06
SALALUGAZI (SA)	11.00	4.00	23.00	2.00	0.08
NASSABA (NS)	14.00	12.00	32.00	3.00	0.11
NFUUKA (NF)	28.00	21.00	58.00	4.00	0.12
KAYINJA (KA)	21.00	12.00	39.00	1.00	0.12
NAKYETENGU (NY)	12.00	13.00	33.00	3.00	0.13
KISUBI (KI)	12.00	11.00	21.00	2.00	0.20
NAMWEZI (NW)	10.00	9.00	33.00	1.00	0.23

Table 5.12 Correspondence analysis of residual nematode scores by cultivars

***** Correspondence analysis *****

*** Squared singular values ***

	CA_Roots	%Roots	Cum%Root
1	0.03278	66.76	66.76
2	0.00888	18.09	84.84
3	0.00744	15.16	100.00

*** Row Scores and Inertias ***

rowsc

Pdata['rows']	1	2	3
1	-0.2856	0.0548	0.1261
2	-0.1035	-0.0006	-0.0333
3	-0.1377	-0.0396	0.0895
4	-0.2784	0.0381	-0.1003
5	-0.1435	-0.1963	-0.1297
6	-0.1391	0.0515	0.1178
7	0.0795	0.0605	0.0337
8	0.1051	-0.0887	-0.1228
9	0.0287	-0.2532	0.1357
10	0.0284	0.0857	-0.0901
11	0.0688	-0.0073	-0.0416
12	0.1735	-0.1364	0.1488
13	0.1842	0.0916	0.0281
14	0.1816	0.0367	-0.0178
15	0.1943	-0.0807	-0.0729
16	0.2372	0.1434	0.0307
17	0.0986	0.1543	-0.0866
18	0.3784	-0.0079	0.0226

R_Inrtia

Pdata['rows']	1	2	3
1	0.0053700	0.0001978	0.0010470
2	0.0005553	0.0000000	0.0000573
3	0.0010054	0.0000831	0.0004247
4	0.0089715	0.0001678	0.0011657
5	0.0009290	0.0017378	0.0007588
6	0.0012495	0.0001715	0.0008958
7	0.0007321	0.0004232	0.0001315
8	0.0005111	0.0003641	0.0006985
9	0.0000392	0.0030479	0.0008750
10	0.0000113	0.0001030	0.0001138
11	0.0005169	0.0000058	0.0001884
12	0.0007337	0.0004533	0.0005397
13	0.0012616	0.0003122	0.0000293
14	0.0022300	0.0000910	0.0000213
15	0.0016792	0.0002898	0.0002364
16	0.0020920	0.0007641	0.0000351
17	0.0002727	0.0006677	0.0002103
18	0.0046239	0.0000020	0.0000164

*** Column Scores and Inertias ***

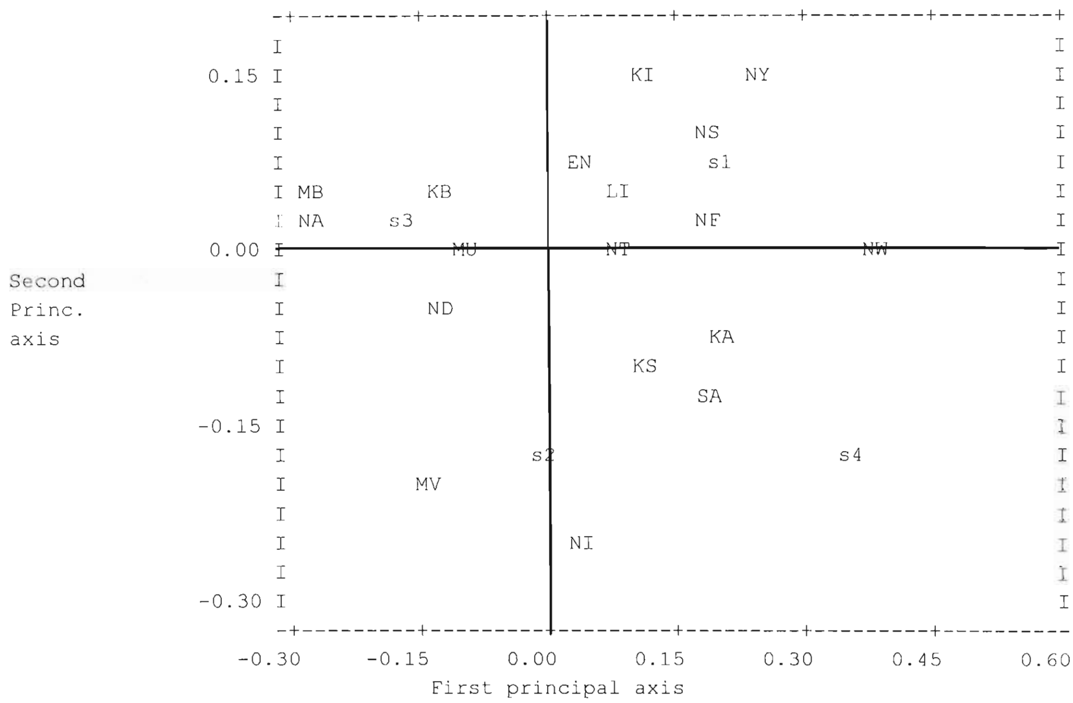
colsc

Pdata['columns']	1	2	3
1	-0.1968	-0.0776	-0.0419
2	0.0161	0.1717	-0.1061
3	0.1756	-0.0295	0.0362
4	-0.3409	0.1660	0.2834

C_Inrtia

Pdata['columns']	1	2	3
1	0.012512	0.001947	0.000567
2	0.000045	0.005047	0.001929
3	0.013924	0.000393	0.000591
4	0.006303	0.001495	0.004357

Figure 5.2 Correspondence analysis of Nematode data



5.4 Measuring leaf spot damage.

5.4.1 Data organization

A data set having 2219 sampling units was created from averages of plants of the same cultivar per farm within a sampling period. The variable YLS(Youngest leaf spot) was not subjected to any transformation. Since YLS is the number of the youngest diseased leaf, the value YLS=1 would be the worst case implying very high susceptibility and YLS=14 would imply high resistance. During the data collection exercise, the plant sampled showing no leaf spot infestation was given a score of 0. However it is not reasonable to allocate a score of 0 to no infestations when the higher the infestations score the lower the actual disease. Two approaches were taken. Firstly zeros were changed to 15 and the data analysed. Secondly all the zeros were excluded from the data leaving 1199 observations and these were separately analysed.

5.4.2 Analysis on data when no diseases is scored 15.

5.4.2.1 Mixed model analyses.

A mixed model analysis was carried out on the above data set with elevation, rainfall, Prainfall and cultivar as fixed effects. Table 5.13 shows that these fixed effects all affect leaf spot infestation. Higher elevation and lower rainfall indicates a lower infestation. Among the cultivars Kayinja, Kisubi and Ndiizi seems to be more resistant to leaf spot infestation, while Salalugazi, Kibuzi and Nakyetengu are more susceptible in relation to the other cultivars. All these effects show statistical significances.

5.4.2.2 Residual Analysis

To examine cultivar differences further, a model (table 5.14) with cultivar as a random effect was fitted with elevation and both rainfalls as fixed effects. The residuals whose effects of elevation and rainfall have been removed where saved for the next analysis.

Univariate analysis was carried out on the residuals so as to subdivide the score into four categories using the quartiles. Table 5.15 shows the output with residual scores showing a strange bimodal distribution, this mainly due to having many of the zero scores brought into the distribution.

A frequency table was created using these categories and cross tabulated by cultivar (see table 5.16)

The results of the tabulated correspondence analysis of frequencies is given table 5.17 and figure 5.3

The first squared singular value explains 74% of the total variation, showing high explanation of the results by one axis alone. The row inertias for cultivars are dominated by cultivars with extreme levels of infestations, Salalugazi and Namwezi with high susceptibility to leaf spot attack and Kayinja which shows high resistance to leaf spot infestations. On the column inertias , category 1 (high infestation) and category 4(low infestation) dominate the other categories and hence suitable for discriminatory purposes.

Table 5.13 Reml analysis of fixed effects of elevation, rainfall and cultivar on leaf spot infestation

```
***** REML Variance Components Analysis *****

Response Variate : yls

Fixed model      : Constant+elev+rain+prainfall+cultivar
Random model     : site+site.farm+site.farm.cultivar

Number of units  : 2219
Absorbing factor : site

* Residual term has been added to model
* All covariates centred

*** Convergence monitoring ***

Cycle      Deviance  Current variance parameters: gammas, sigma2, others
    0      *        1.00000   1.00000   1.00000   1.00000
    1    13100.1   1.01230   0.159987 0.0968873   6.53712
    2    6780.35  0.995005  0.155114 0.0227856   6.71270
    3    6768.73  0.993362  0.152939 0.0105796   6.76829
    4    6768.27  0.992876  0.152254 0.83266E-02   6.78033
    5    6768.25  0.992763  0.152104 0.78951E-02   6.78272


*** Estimated Variance Components ***

Random term                Component        S.e.

site                           6.734       2.121
site.farm                      1.032       0.217
site.farm.cultivar             0.054       0.129
*units*                        6.783       0.230
                   Deviance   d.f.

                    6798.05  2195

*** Wald tests for fixed effects ***
    Fixed term              Wald statistic      d.f.

   Elev                          29.3             1
   Rainfall                       2.8             1
   Prainfall                     58.2             1
   cultivar                     331.1            17
```

*** Table of effects

Effects	Estimates	stderr
Constant	9.654	0.8374
Elevation	0.01304	0.0022
Rainfall	-0.0051	0.0009
Prainfall	0.01078	0.0014

*** Table of predicted means for cultivars

Cultivar	Estimates
Kayinja	14.192 a
Kisubi	12.529 ab
Ndiizi	10.888 ab
Ensenyi	9.654 b
Kisansa	9.529 b
Musakala	9.489 b
Nfuuka	9.449 b
Likhago	9.359 b
Namwezi	9.331 b
Nakabululu	9.291 b
Muvubo	9.156 b
Nassaba	9.177 b
Mbwazirume	8.995 b
Ndiibwabalangira	8.876 b
Kibuzi	8.704 b
Nakyetengu	8.695 b
Salalugazi	8.486 b

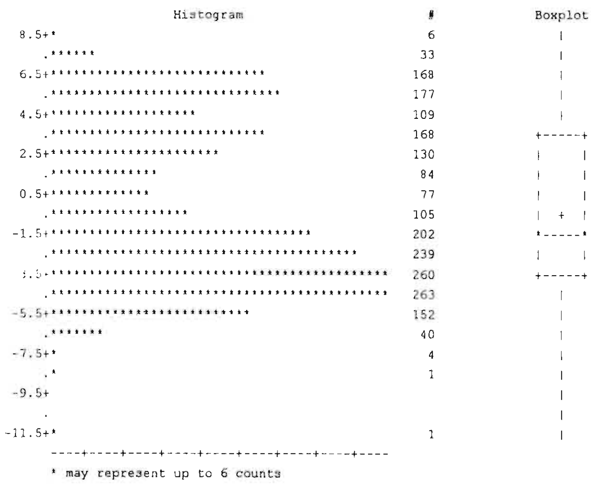
Standard error of differences: Average 0.3064

Table 5.14 Reml analysis on leaf spot scores to derive residuals.

```
***** REML Variance Components Analysis *****


Response Variate : yls

Fixed model      : Constant+elev+rain+prainfall
Random model     : site+site.farm+site.farm.cultivar


Number of units  : 2219
Absorbing factor : site

* Residual term has been added to model
* All covariates centred



*** Convergence monitoring ***

Cycle    Deviance  Current variance parameters: gammas, sigma2, others
    0       *       1.00000   1.00000   1.00000   1.00000
    1    13475.0  0.963014  0.264935  0.249103   6.53057
    2    6975.07  0.933331  0.274036  0.206407   6.60736
    3    6973.57  0.930686  0.279924  0.200087   6.62044
    4    6973.52  0.930259  0.281350  0.198918   6.62275


*** Estimated Variance Components ***

Random term                Component        S.e.

site                           6.161       2.021
site.farm                      1.863       0.386
site.farm.cultivar             1.317       0.218
*units*                        6.623       0.227


*** Deviance: -2*Log-Likelihood ***


                   Deviance   d.f.

                    6975.50  2212

Note: deviance omits constants which depend on fixed model fitted.



*** Wald tests for fixed effects ***


   Fixed term               Wald statistic      d.f.

   elev                           27.6             1
   rainfall                        3.2             1
   prainfall                      58.2             1

* All Wald statistics are calculated ignoring terms fitted later in the model
```

Table 5.15 Distribution of residual scores for leaf spot analysis.

```
                              Univariate Procedure

Variable=YRESID

                                    Moments
                  N             2219  Sum Wgts       2219
                  Mean      -0.12423  Sum        -275.656
                  Std Dev   4.068277  Variance   16.55088
                  Skewness  0.304206  Kurtosis   -1.26568
                  USS       36744.09  CSS        36709.84
                  CV        -3274.92  Std Mean   0.086364
                  T:Mean=0   -1.4384  Pr>|T|       0.1505
                  Num ^= 0      2219  Num > 0         952
                  M(Sign)     -157.5  Pr>=|M|      0.0001
                  Sgn Rank  -22007.5  Pr>=|S|      0.4661
                              Quantiles(Def=5)
                   100% Max    8.383       99%     7.197
                    75% Q3     3.657       95%     6.398
                    50% Med   -1.372       90%     5.916
                    25% Q1    -3.643       10%    -4.878
                     0% Min  -11.047        5%    -5.482
                                            1%    -6.214

                   Range       19.43
                   Q3-Q1         7.3
                   Mode        4.485
                                  Extremes

                     Lowest    Obs      Highest    Obs
                    -11.047(   1482)    8.177(     276)
                      -8.79(   1478)     8.22(    1332)
                     -7.372(   2068)     8.22(    1336)
                     -7.369(   1008)    8.304(     274)
                     -7.115(   1004)    8.383(     565)
                             Univariate Procedure

Variable=YRESID
```



Variable=YRESID

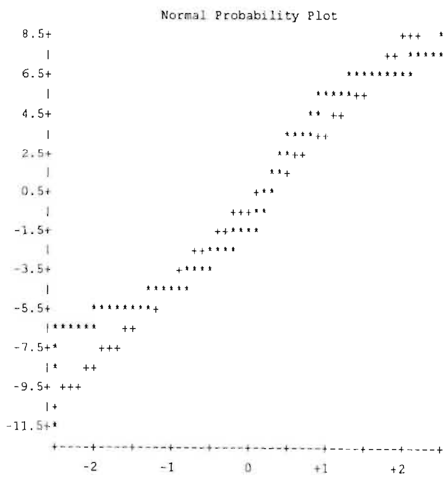


Table 5.16 Residual scores for leaf spot infestation by cultivar

	Residual scores for leaf spot				
	<-3.6	>-3.6&<-1.3	>-1.3&<3.6	>3.6&<8.4	
Cultivars	Frequency	Frequency	Frequency	Frequency	MEAN
SALALUGAZI (SA)	33	13	4	3	-3.61
NAMWEZI (NW)	40	29	6	6	-2.90
NASSABA (NS)	25	21	18	4	-1.93
NDIIBWABALANGIRA(NI)	37	50	26	9	-1.82
NFUUKA (NF)	40	40	24	12	-1.72
KISANSA (KS)	35	25	18	12	-1.62
NAKYETENGU (NY)	28	23	31	3	-1.60
ENSENYI (EN)	1	15	12	.	-1.43
MUVUBO (MV)	40	30	22	24	-0.87
MUSAKALA (MU)	45	44	24	45	-0.27
MBWAZIRUME (MB)	33	39	45	37	-0.19
KIBUZI (KB)	44	28	29	49	0.24
NAKITEMBE (NT)	48	60	54	72	0.26
LIKHAGO (LI)	50	48	75	63	0.44
NAKABULULU (NA)	29	51	70	64	0.94
NDIIZI (ND)	21	20	62	59	1.69
KISUBI (KI)	4	6	12	27	2.69
KAYINJA (KA)	2	13	24	64	3.81

Table 5.17 Correspondence analysis of residual leaf spot scores by cultivars

*** Squared singular values ***

	CA_Roots	%Roots	Cum%Root
1	0.14001	73.95	73.95
2	0.03604	19.03	92.98
3	0.01329	7.02	100.00

*** Row Scores and Inertias ***

rowsc

Pdata['rows']	1	2	3
1	0.7813	-0.4117	0.2151
2	0.6801	-0.2522	-0.0725
3	0.4410	0.1215	0.0774
4	0.4325	0.1471	-0.1836
5	0.3989	0.0380	-0.0634
6	0.3669	-0.0757	0.0614
7	0.3623	0.2999	0.2032
8	0.2154	0.8331	-0.3810
9	0.2025	-0.1289	0.0227
10	0.0504	-0.1897	-0.1159
11	-0.0448	0.1018	0.0062
12	-0.0769	-0.2267	0.0582
13	-0.1234	-0.0382	-0.0828
14	-0.1292	0.0889	0.0964
15	-0.2453	0.1605	-0.0221
16	-0.4438	0.1228	0.1815
17	-0.7080	-0.2036	-0.0392
18	-0.8734	-0.2272	-0.1288

R_Inrtia

Pdata['rows']	1	2	3
1	0.014580	0.004048	0.001105
2	0.016886	0.002321	0.000192
3	0.005960	0.000453	0.000184
4	0.010284	0.001189	0.001854
5	0.008319	0.000075	0.000210
6	0.005461	0.000232	0.000153
7	0.005027	0.003444	0.001581
8	0.000585	0.008757	0.001832
9	0.002144	0.000869	0.000027
10	0.000181	0.002563	0.000956
11	0.000139	0.000720	0.000003
12	0.000400	0.003475	0.000229
13	0.001605	0.000154	0.000724
14	0.001774	0.000840	0.000989
15	0.005803	0.002484	0.000047
16	0.014376	0.001101	0.002405
17	0.011070	0.000916	0.000034
18	0.035412	0.002397	0.000770

*** Column Scores and Inertias ***

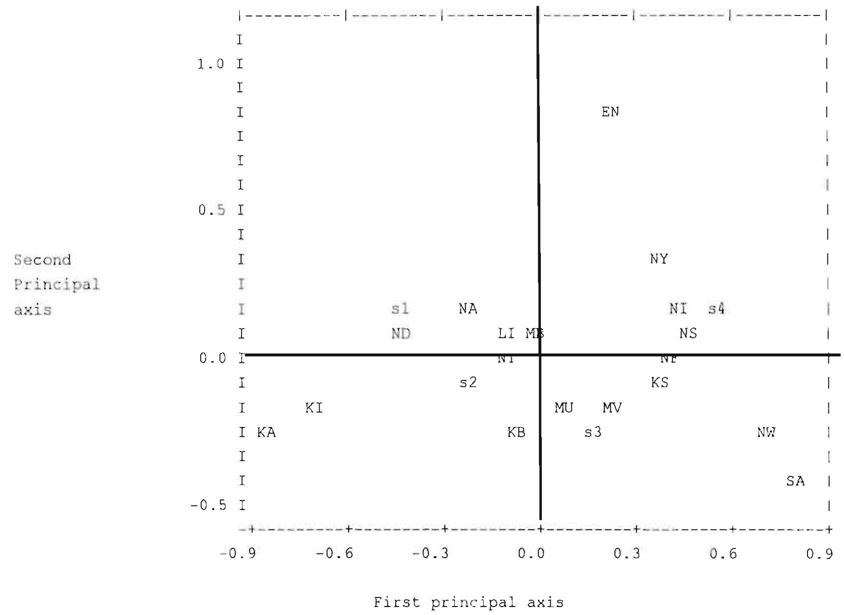
colsc

Pdata['columns']	1	2	3
1	0.4396	-0.1787	0.0986
2	0.2468	0.0959	-0.1751
3	-0.1571	0.2640	0.1082
4	-0.5309	-0.1823	-0.0320

C_Inrtia

Pdata['columns']	1	2	3
1	0.04833	0.00799	0.00243
2	0.01524	0.00230	0.00767
3	0.00619	0.01746	0.00293
4	0.07025	0.00828	0.00025

Figure 5.3 Correspondence analysis of Leafspot data



5.4.3 Analysis on data excluding zero scores

As explained in the earlier section of this chapter, a dataset of 1393 observations was created after removing all the zero scores for the variable youngest leaf spotted with disease (yls), in this case the lowest score was 1 and the highest 14.

5.4.3.1 Mixed model analyses.

A mixed model analyses was carried out on the above data set with elevation, both rainfalls and cultivar as fixed effects. Table 5.18 shows that these fixed effects all affect leaf spot infestation in the same way as the previous analyses when the zero's are included. Among the cultivars Kayinja and Kisubi seems to be more resistant to leaf spot infestation, while Mbwazirume and Nakyetengu display high susceptibility in relation to the other cultivars. The results show some differences with cultivar susceptibility and resistance with data when zeros are included.

5.4.3.2 Residual Analysis

To examine cultivar differences further, a model (table 5.19) with cultivar as a random effect was fitted with elevation and both rainfall as fixed effects. The residuals whose effects of elevation and rainfall have been removed were saved for the next analysis.

Univariate analysis was carried out on the residuals so as to subdivide the score into four categories using the quartiles. Table 5.20 shows the output with residual scores showing a right skewed distribution pattern but could be termed as fairly normal.

Table 5.21 shows frequency distribution of the categorised variable into quartiles and hence correspondence analysis performed on the frequencies is shown in table 5.22 and figure 5.4

The first squared singular value explain over 80% of the total variation, showing high explanation of the results by one axis alone. The row inertias for cultivars are dominated by cultivars with extreme levels of infestations, Salalugazi with high susceptibility to leaf spot attack and Kayinja which shows high resistance to leaf spot infestations. On the column inertias , category 1 (high infestation) and category 4(low infestation) dominate the other categories and hence suitable for discriminatory purposes.

Table 5.18 Reml analysis of fixed effects of elevation, rainfall and cultivar on leaf spot infestation

```
***** REML Variance Components Analysis *****
Response Variate : yls
Fixed model      : Constant+elev+rain+prainfall+cultivar
Random model     : site+site.farm+site.farm.cultivar

Number of units  : 1393
Absorbing factor : site
* Residual term has been added to model
* All covariates centred
* Analysis is subject to the restriction on yls

*** Convergence monitoring ***

Cycle      Deviance  Current variance parameters: gammas, sigma2, others
    0         *       1.00000   1.00000   1.00000   1.00000
    1      5130.47  0.954769 0.0574095 0.0784034   3.93546
    2      3561.29   1.01083 0.0622281 0.0431020   3.99213
    3      3559.79   1.01781 0.0641364 0.0372241   4.00495
    4      3559.74   1.01862 0.0646512 0.0361249   4.00731
    5      3559.74   1.01873 0.0647743 0.0359117   4.00774

*** Estimated Variance Components ***

Random term              Component        S.e.

site                         4.083       1.314
site.farm                    0.260       0.104
site.farm.cultivars          0.144       0.114
*units*                      4.008       0.176
*** Deviance: -2*Log-Likelihood ***
                   Deviance   d.f.
                    3561.97  1369

*** Wald tests for fixed effects ***
   Fixed term              Wald statistic      d.f.

   Elvation                        18.1           1
   Rainfall                        10.9           1
   Prainfall                       55.4           1
   Cultivars                      215.1          17

Table of effects

    Effects           Estimates     Stderr
    Constant          7.952          1.092
    Elevation         0.009          0.001
    Rainfall          0.009          0.001
    Prainfall         0.019          0.001

Tablw of predcited means
     Cultivar              Estimates
     Kayinja               11.476
     Kisubi                 9.744
     Ndiizi                 8.099
     Nassba                 8.036
     Ensenyi                7.952
     Nfuuka                 7.681
     Nakitembe              7.445
     Kisansa                7.431
     Likhago                7.389
     Musakala               7.128
     Nakabululu             7.097
     Namwezi                7.006
     Salalugazi             6.716
     Ndiibwabalangira       6.684
     Muvubo                 6.676
     Kibuzi                 6.555
     Mbwazirume             6.524
     Makyetengu             6.418

     Standard error of difference: Average 0.4695
```

Table 5.19 Reml analysis on leaf spot scores to derive residuals.

```
***** REML Variance Components Analysis *****


Response Variate : yls

Fixed model      : Constant+elev+rain+prain
Random model     : site+site.farm+site.farm.cv


Number of units  : 1393
Absorbing factor : site

* Residual term has been added to model
* All covariates centred

* Analysis is subject to the restriction on yls


*** Convergence monitoring ***

Cycle      Deviance  Current variance parameters: gammas, sigma2, others
    0         *        1.00000   1.00000   1.00000   1.00000
    1      5304.75  0.795495  0.147436  0.248825   3.94504
    2      3698.55  0.800768  0.132635  0.212390   4.00128
    3      3697.73  0.801900  0.132069  0.206458   4.01250
    4      3697.71  0.802018  0.132540  0.205166   4.01456
    5      3697.71  0.802008  0.132836  0.204805   4.01501


*** Estimated Variance Components ***

Random term                Component        S.e.

site                           3.220       1.085
site.farm                      0.533       0.183
site.farm.cv                   0.822       0.173
*units*                        4.015       0.179


*** Deviance: -2*Log-Likelihood ***


                   Deviance   d.f.

                    3698.08  1386

Note: deviance omits constants which depend on fixed model fitted.



*** Wald tests for fixed effects ***


   Fixed term              Wald statistic      d.f.

   Elevation                       22.5            1
   Rainfall                         9.0            1
   Prainfall                       52.1            1

  Table of effects

       Effects      Estimates      Stderr
       Constant       7.337        0.417
       Elevation      0.009        0.001
       Rainfall       0.001        0.001
       Prainfall      0.010        0.001
```

Table 5.20 Distribution of residual scores for leaf spot analysis.

```
                                   Univariate Procedure

Variable=YRESID

                                         Moments

                      N              1393  Sum Wgts        1393
                      Mean       -0.92245  Sum         -1284.97
                      Std Dev    2.661939  Variance    7.085919
                      Skewness   1.028903  Kurtosis     0.40697
                      USS        11048.92  CSS           9863.6
                      CV         -288.573  Std Mean    0.071322
                      T:Mean=0   -12.9336  Pr>|T|        0.0001
                      Num ^= 0       1393  Num > 0          369
                      M(Sign)      -327.5  Pr>=|M|       0.0001
                      Sgn Rank    -202973  Pr>=|S|       0.0001

                                    Quantiles(Def=5)

                       100% Max       7.557       99%       6.307
                        75% Q3        0.259       95%       4.883
                        50% Med       -1.64       90%       3.464
                        25% Q1        -2.78       10%      -3.641
                         0% Min      -7.324        5%      -4.126
                                                   1%      -4.646

                       Range         14.881
                       Q3-Q1          3.039
                       Mode           -3.61

                                        Extremes

                          Lowest    Obs     Highest    Obs
                          -7.324(    932)    6.918(    1238)
                          -5.541(    624)    6.945(    1284)
                          -5.503(    527)    7.196(    1166)
                           -5.46(   1267)    7.221(     761)
                          -5.359(    620)    7.557(     804)

                    Histogram                          #             Boxplot
       7.5+*                                           3                0
          .***                                        20                0
          .******                                     39                0
       4.5+*******                                    49                0
          .*********                                  63                |
          .*******                                    47                |
       1.5+*********                                  58                |
          .*************                              90             +-----+
          .*************************                 170             |  +  |
      -1.5+***************************************   269             *-----*
          .*******************************************297             +-----+
          .*****************************             197                |
      -4.5+************                               83                |
          .*                                           7                |
          .                                                             |
      -7.5+*                                           1                |
           ----+----+----+----+----+----+----+----+---
           * may represent up to 7 counts

                                    The SAS System                          2
                                             18:31 Wednesday, February 24, 1999

                                   Univariate Procedure

Variable=YRESID

                                   Normal Probability Plot
                 7.5+                                                  *
                    |                                              *****
                    |                                          ***** +++
                 4.5+                                         **** ++++
                    |                                       ***++++
                    |                                     **+++
                 1.5+                                    ***
                    |                                ++***
                    |                             +++****
                -1.5+                          ++*******
                    |                       *******
                    |                ********
                -4.5+************+++
                    |*      ++++
                    |    +++
                -7.5+*++
                    +----+----+----+----+----+----+----+----+----+----+
                        -2        -1         0        +1        +2
```

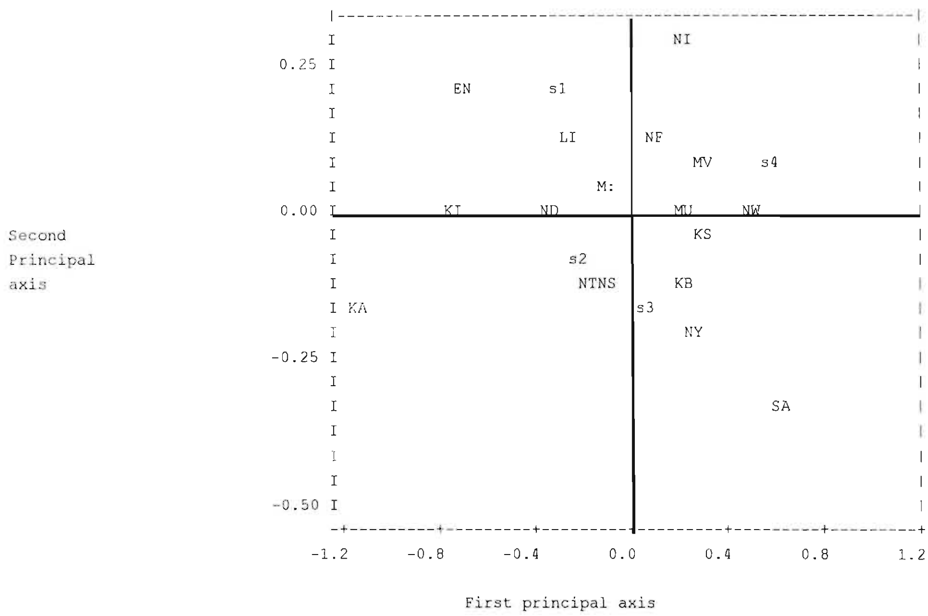
Table 5.21 Residual scores for leaf infestation by cultivar

	Residual scores for leaf spot				
	<-2.7	>-2.7&<-1.6	>-1.6&<0.2	>0.2&7.2	
Cultivars	Frequency	Frequency	Frequency	Frequency	MEAN
SALALUGAZI(SA)	25.00	16.00	6.00	3.00	-2.59
NAMWEZI (NW)	28.00	24.00	20.00	4.00	-2.09
MUVUBO (MV)	21.00	30.00	18.00	10.00	-1.71
NAKYETENGU (NY)	22.00	13.00	12.00	10.00	-1.65
NDIIBWABALANGIRA(NI)	21.00	33.00	36.00	13.00	-1.53
KISANSA (KS)	24.00	19.00	18.00	9.00	-1.48
KIBUZI (KB)	30.00	26.00	17.00	16.00	-1.41
MUSAKALA (MU)	27.00	28.00	23.00	16.00	-1.41
NFUUKA (NF)	23.00	28.00	31.00	18.00	-1.23
MBWAZIRUME (MB)	20.00	17.00	27.00	23.00	-0.85
NAKABULU (NA)	20.00	27.00	24.00	26.00	-0.65
NASSABA (NS)	17.00	14.00	15.00	20.00	-0.63
LIKHAGO (LI)	20.00	32.00	42.00	46.00	-0.30
NAKITEMBE (NT)	38.00	23.00	38.00	48.00	-0.30
NDIIZI (ND)	11.00	11.00	18.00	25.00	-0.28
ENSENYI (EN)	0.00	1.00	2.00	3.00	0.02
KISUBI (KI)	2.00	2.00	8.00	13.00	1.11
KAYINJA (KA)	0.00	4.00	7.00	31.00	3.11

Table 5.22 Correspondence analysis of residual leaf spot scores by cultivars

```
***** Correspondence analysis *****

*** Squared singular values ***
       CA_Roots    %Roots    Cum%Root

  1    0.11641     80.85     80.85
  2    0.02035     14.13     94.98
  3    0.00722      5.02    100.00

*** Row Scores and Inertias ***

          rowsc
              1          2          3
Pdata['rows']
  1      0.5980    -0.3409     0.0658
  2      0.4661     0.0044    -0.0419
  3      0.2927     0.0962     0.1760
  4      0.2265    -0.2133    -0.0703
  5      0.1809     0.2831    -0.0213
  6      0.2869    -0.0441    -0.0569
  7      0.2178    -0.1319     0.0740
  8      0.1841     0.0072     0.0419
  9      0.0941     0.1426    -0.0270
 10     -0.1025     0.0362    -0.1186
 11     -0.0686     0.0458     0.0908
 12     -0.1197    -0.1110     0.0039
 13     -0.2684     0.1224     0.0130
 14     -0.1985    -0.1263    -0.1057
 15     -0.3732    -0.0129    -0.0391
 16     -0.7335     0.2066     0.0249
 17     -0.7471     0.0155    -0.1208
 18     -1.1571    -0.1815     0.2083

       R_Inrtia
              1          2          3
Pdata['rows']
  1    0.012835   0.004172   0.000156
  2    0.011851   0.000001   0.000096
  3    0.004858   0.000524   0.001757
  4    0.002099   0.001862   0.000202
  5    0.002420   0.005926   0.000034
  6    0.004136   0.000098   0.000163
  7    0.003032   0.001111   0.000350
  8    0.002286   0.000004   0.000119
  9    0.000635   0.001459   0.000052
 10    0.000656   0.000082   0.000878
 11    0.000327   0.000146   0.000574
 12    0.000679   0.000584   0.000001
 13    0.007238   0.001507   0.000017
 14    0.004159   0.001683   0.001180
 15    0.006501   0.000008   0.000071
 16    0.002318   0.000184   0.000003
 17    0.010016   0.000004   0.000262
 18    0.040370   0.000994   0.001309

*** Column Scores and Inertias ***

          colsc
              1          2          3
Pdata['columns']
  1      0.3318    -0.1884    -0.0467
  2      0.2355     0.0949     0.1227
  3     -0.0313     0.1693    -0.1016
  4     -0.5582    -0.0856     0.0311


       C_Inrtia
              1          2          3
Pdata['columns']
  1     0.02759    0.00889    0.00055
  2     0.01386    0.00225    0.00376
  3     0.00025    0.00745    0.00268
  4     0.07472    0.00176    0.00023
```

Figure 5.4 Correspondence analysis of Leafspot data



CHAPTER 6

Mixed Model Analysis

As discussed in Chapter 3, Mixed model analysis in this chapter has been used to explain infestation in relation to environmental and farm management factors for weevils, root knot nematodes and leaf spot diseases. Error structures due to site, farm within site have been estimated as random effects while errors due to cultivars within farm have been estimated as both random and fixed effects. Various important fixed effects are then fitted and tested for significance of their contribution to the model.

6.1 REML on Weevils infestation

Chapter 5 discussed the use of first principal components of the weevil infestation variables as being a simple arithmetic average over all infestation variables. In this chapter, two approaches of mixed model analysis have been performed. The first one is using the first principal component score as a measure of weevil infestation and subjecting it to various models with different error structures. We then carried out analysis using the average total damage variable (xt) but transformed into a log scale, for the purpose of comparing the results.

6.1.1 First Principal component score as dependent variable for weevils infestation

The first principal component score derived from the PCA on the weevil infestation log transformed variables, as in chapter 5, was subjected to mixed model analysis so as to establish and quantify the relationships of weevil infestation with environmental and farm management practices. Several models were fitted and using the significance of Wald's statistic, variables were dropped which showed no significance ; the significant ones were re-fitted until the final model was achieved. Table 6.1 shows the results of the analysis when using site, farm within site and cultivars within farms as random effects while table 6.3 shows results after fitting site, farm within

site as random effect and cultivars now treated as fixed effects only.
The final fitted model (Model 3*) in table 6.1 does suggest that elevation, rainfall, previous month rainfall, cultivars, Calcium, Sodium and Nitrogen to be the major factors affecting weevil infestation. Among the cultivars Salalugazi and Nassaba seem to be highly susceptible to weevil attack while Kisubi most resistant followed by Kayinja and Ndiizi . The coefficients of these factors indicate that a weevil infestation decreases with increase of elevation, decrease with increase of rainfall, decrease with increase of rainfall prior to the month of sampling, differ from cultivar to cultivar, increase with increase of calcium content in the soil and decrease with increase of Sodium (Na) and Nitrogen(N) content in the soil.

The same trend of effects is shown in Table 6.3 where the random component only focuses on site, farm within site variance component. The final fitted model (Model 3 *) shows the same results as in table 6.1 confirming the same fixed effects to be still very significant on the weevil infestation. Table 6.4 shows Salalugazi and Nassaba being the most susceptibility while Kisubi, Kayinja and Ndiizi show high resistant to weevil attack. The results are similar to results of table 6.2 when fitting cultivars as both random and fixed effects.

The coefficient of the final model carry the same signs and are almost equal in magnitude. In this case treating cultivars as fixed or random effects yields the same results.
Models 4 in table 6.1 and 6.3 illustrate the colinearity effect of Calcium (Ca) with other soil nutrients (i.e Sodium and Nitrogen). When Calcium is dropped from the model, Sodium and Nitrogen are no longer significant as shown by the Wald's statistic in model 4. A high Calcium saturation indicates a favorable pH for plant growth and microbial activity. Also, a prominence of Ca will usually mean low concentrations of undesirable exchangeable cations such as Al^{3+} in acidic soils and Na^{+} in sodic soils (Barber, S.A 1984; Follet, R. H., Murphy, L. S, Donahue, R. L 1981). Many crops will respond to Ca applications when the %Ca^{2+} saturation falls below 25% (Mortvedt, J. J., and Fox F.R (1985).

While Ca^{2+} uptake is depressed by NH_4^+, K^+,Al^{3+}, its absorption is increased when plants are supplied with NO_3^-. A high level of NO_3^- nutrition stimulates organic anion synthesis and the resultant accumulation of cations, particularly Ca^{2+} (Beaton J.D., Fox R. L, and Jones, M. B. (1985)). Plants absorb N as both NH4+ and NO3-. NO3- generally occurs in higher concentrations than NH4-, and it is free to move to the roots by mass flow and diffusion. Some NH4+ is always present and will influence plant growth and metabolism in ways that are not completely understood (Tisdale S. L, Nelson, W. L, Beaton, J. D., Havlin, J. L (1985)). Preference of plants for either NH4+ or NO3- is determined by the age and type of plant, the environment, and other factors (Power, J. F., and Papendick R. I. (1985).

Table 6.1 First principal score analysis with Cultivars as both fixed and random effects for weevil infestation

No. Of observation=1470

Model 1		Model 2		Model 3*		Model 4	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	0.649	Site	0.683	Site	0.695	Site	0.650
Site.farm	0.454	Site.farm	0.396	Site.farm	0.397	Site.farm	0.452
Site.farm.cv	0.220	Site.farm.cv	0.214	Site.farm.cv	0.213	Site.farm.cv	0.215
Units	1.794	Units	1.796	Units	1.796	Units	1.795
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term*	Wald Statistic	Fixed term	Wald Statistic
Elevation	20.0	Elevation	18.9	Elevation	18.6	Elevation	20.0
Rainfall	22.8	Rainfall	22.8	Rainfall	22.8	Rainfall	22.8
Prainfall	7.2	Prainfall	7.1	Prainfall	7.1	Prainfall	7.2
Cultivars	255.2	Cultivars	257.7	Cultivars	257.9	Cultivars	256.5
Ca	1.8	Ca	2.0	Ca	2.0	Na	0.1
pH	1.1	Na	2.9	Na	2.9	N	0.2
Intercrop	0.2	N	4.1	N	4.1		
Mulching	0.0	pH	0.9				
Weeding	0.5						
De-leafing	0.1						
Desuckering	1.0						
OM	0.0						
N	2.7						
K	0.4						
Na	3.1						
Sand	0.5						
Clay	0.0						
Silt	0.1						
Mg	0.6						

Model 3* The Final model:
Y=3.113-0.003591 Elevation-0.004969 Rainfall-0.002449 Prainfall+Cultivars+0.08760 Ca-1.848 Na-2.165N

Codes of importance to note : OM - Organic Matter, cv - Cultivar , Prainfall - Rainfall prior to the month of sampling.

Table 6.2 Table of predicted means for cultivars using PCA on weevil infestation when cultivars are fitted as both random and fixed effects

Cultivars	Predicted Means for principal score
Salalugazi	4.529
Nassaba	4.373
Kibuzi	3.855
Kisansa	3.746
Namwezi	3.744
Likhago	3.713
Nakabululu	3.708
Nakitembe	3.705
Mbwazirume	3.674
Nfuuka	3.643
Musakala	3.577
Ndiibwabalangira	3.472
Muvubo	3.463
Ensenyi	3.113
Nakyetengu	3.023
Ndiizi	1.734
Kayinja	1.493
Kisubi	0.899

Standard error of differences: 0.3290

Table 6.3 First principal component score with cultivars as fixed effect on weevil infestation.

No. Of observation=1470

Model 1		Model 2		Model 3*		Model 4	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	0.623	Site	0.657	Site	0.669	Site	0.624
Site.farm	0.491	Site.farm	0.432	Site.farm	0.431	Site.farm	0.487
Units	1.956	Units	1.956	Units	1.955	Units	1.955
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic
Elevation	21.2	Elevation	20.0	Elevation	19.7	Elevation	21.2
Rainfall	21.0	Rainfall	21.0	Rainfall	21.0	Rainfall	21.0
Prainfall	6.4	Prainfall	6.4	Prainfall	6.4	Prainfall	6.4
Cultivars	324.7	Cultivars	326.1	Cultivars	326.2	Cultivars	324.9
Ca	1.6	Ca	1.8	Ca	1.8	Na	0.2
pH	1.1	Na	3.1	Na	3.1	N	0.3
Intercrop	0.2	N	4.5	N	4.5		
Mulching	0.0	pH	0.8				
Weeding	0.6						
De-leafing	0.1						
Desuckering	0.9						
OM	0.1						
N	2.9						
K	0.5						
Na	3.2						
Sand	0.4						
Clay	0.0						
Silt	0.1						
Mg	0.6						

* Final model:

Y=3.105-0.003678 Elevation-0.004926 Rainfall-0.002319 Prainfall+Cultivars+0.08720 Ca-1.870 Na-2.225 N

Important codes: OM - Organic matter, Prainfall - Rainfall prior to the month of sampling.

Table 6.4 Table of predicted means for cultivars using PCA on weevil infestation when cultivars are fitted as fixed effect only

Cultivars	Predicted Means for principal score
Salaluga	4.609
Nassaba	4.393
Kibuzi	3.847
Kisansa	3.781
Namwezi	3.768
Likhago	3.736
Nakabululu	3.747
Nakitembe	3.726
Mbwazirume	3.666
Nfuuka	3.652
Musakala	3.601
Ndiibwabi	3.462
Muvubo	3.490
Ensenyi	3.105
Nakyetengu	3.052
Ndiizi	1.720
Kayinja	1.447
Kisubi	0.893

Standard error of differences: 0.2947

6.1.2 Using the log(xt+1) as dependent variable for weevil infestation

The variable xt(average damage) derived by either averaging inner crossection (xi) and outer crossection(xo), or averaging lower crossection damage (xl) and upper crossection damage(xu) could well be used as a principal measure of the weevil infestation. Results in table 6.5 and 6.7 do give similar results to table 6.1 and 6.3, for the principal score analyses. The final fitted models indicated by(model 3 *) in both tables 6.5 and 6.7 include the same variables as those of final models in table 6.1 and 6.3, the coefficients have also consistent signs. Again models 4 in table 6.5 and 6.7 demonstrates the colinearity effect of dropping calcium to the other soil nutrients (Sodium and Nitrogen). The relationship of these soil nutrients is discussed in section 6.1.1 of this chapter. The similarity of results of variable xt and the principal score implies that xt is as good in measuring the weevil infestation, and may be used as a principal measure of weevil infestation. Table 6.6 and 6.8 shows Salalugazi, Nassaba and Kibuzi as still the most susceptible cultivars while Kisubi, Kayinja and Ndiizi the most resistant to weevil infestation.

Table 6.5 Using Log (xt+1) as dependent variable with Cultivars as both Fixed and Random effects

No. Of observation=1470

Model 1		Model 2		Model 3*		Model 4	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	0.0569	Site	0.0620	Site	0.0638	Site	0.0594
Site.farm	0.0584	Site.farm	0.0503	Site.farm	0.0510	Site.farm	0.0570
Site.farm.cv	0.0249	Site.farm.cv	0.0242	Site.farm.cv	0.0240	Site.farm.cv	0.0244
Units	0.2449	Units	0.2452	Units	0.2452	Units	0.2451
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic
Elevation	19.7	Elevation	18.2	Elevation	17.7	Elevation	18.9
Rainfall	18.9	Rainfall	18.8	Rainfall	18.8	Rainfall	18.8
Prainfall	6.5	Prainfall	6.4	Prainfall	6.5	Prainfall	6.5
Cultivars	227.3	Cultivars	229.4	Cultivars	229.7	Cultivars	228.4
Ca	1.6	Ca	1.9	Ca	1.9	Na	0.0
pH	1.9	Na	2.1	Na	2.1	N	0.3
Intercrop	0.2	N	3.9	N	3.8		
Mulching	0.1	pH	1.6				
Weeding	0.4						
De-leafing	0.1						
Desuckering	0.7						
OM	0.0						
N	2.5						
K	0.6						
Na	2.1						
Sand	0.5						
Clay	0.0						
Silt	0.0						
Mg	0.9						

* **Final model**:

Y=0.8564-0.001098*Elevation-0.001667*Rainfall-0.0008550*Prainfall+cultivars+0.02883*Ca-0.5616*Na-0.7529*N

Important codes: cv - Cultivar, Prainfall - Rainfall prior to the month of sampling, OM - Organic matter

Table 6.6 Table of predicted means for cultivars using log(xt+1) while cultivars are fitted as both random and fixed effects

Cultivars	Predicted Means for log(xt+1)
Salalugazi	1.4346
Kibuzi	1.1408
Nassaba	1.3669
Nakabululu	1.1361
Kisansa	1.1364
Nakitembe	1.1313
Namwezi	1.1199
Likhago	1.1131
Musakala	1.1042
Mbwazirume	1.1034
Nfuuka	1.0771
Ndiibwabalangira	1.0453
Muvubo	1.0252
Nakyetengu	0.8971
Ensenyi	0.8564
Ndiizi	0.4502
Kayinja	0.3499
Kisubi	0.2041

Standard error of differences: 0.1183

Table 6.7 Using log (xt+1) as dependent variable with Cultivars as Fixed effect only

No. Of observation=1470

Model 1		Model 2		Model 3*		Model 4	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	0.0546	Site	0.0596	Site	0.0614	Site	0.0231
Site.farm	0.0627	Site.farm	0.0546	Site.farm	0.0551	Site.farm	0.0126
Units	0.2633	Units	0.2633	Units	0.2631	Units	0.0102
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic
Elevation	21.0	Elevation	19.4	Elevation	18.9	Elevation	20.2
Rainfall	17.6	Rainfall	17.5	Rainfall	17.5	Rainfall	17.5
Prainfall	6.0	Prainfall	6.0	Prainfall	6.0	Prainfall	6.1
Cultivars	281.3	Cultivars	282.6	Cultivars	282.5	Cultivars	281.5
Ca	1.5	Ca	1.7	Ca	1.7	Na	0.1
pH	1.8	Na	2.3	Na	2.2	N	0.4
Intercrop	0.2	N	4.2	N	4.2		
Mulching	0.1	pH	1.5				
Weeding	0.5						
De-leafing	0.2						
Desuckering	0.6						
OM	0.0						
N	2.7						
K	0.7						
Na	2.2						
Sand	0.4						
Clay	0.0						
Silt	0.0						
Mg	0.9						

* Final model

Y=0.8555-0.001129*Elevation-0.001651*Rainfall-0.0008243*Prainfall+Cultivars+0.02889*Ca-0.5730Na-0.7716*N

Important codes: Prainfall - Rainfall prior to the month of sampling, OM - Organic Matter

Table 6.8 Table of predicted means for cultivars using log(xt+1) as response variable while cultivars are fitted as fixed effects

Cultivars	Predicted Means for log(xt+1)
Salalugazi	1.4563
Nassaba	1.3722
Nakabululu	1.1483
Kisansa	1.1454
Kibuzi	1.1378
Nakitembe	1.1373
Namwezi	1.1261
Likhago	1.1216
Musakala	1.1095
Mbwazirume	1.1004
Nfuuka	1.0804
Ndiibwabalangira	1.0415
Muvubo	1.0323
Nakyetengu	0.9021
Ensenyi	0.8555
Ndiizi	0.4470
Kayinja	0.3334
Kisubi	0.1964

Standard error of differences: 0.1077

6.2 REML on Nematode infestation

6.2.1 Using root necrosis on log scale as dependent variable.

Two forms of random effects model was fitted. First, a model with site, farm within site and cultivars within farm as random effects was fitted in REML procedure using log(ri+1) as the dependent variable, and secondly using only sites and farms within sites as random effects. Tables 6.9 and 6.11 gives results of several models fitted using these different error structures and the final fitted model (Model 5 *) obtained using the significance of the Wald's statistic to select or drop terms. Tables 6.10 and 6.12 shows the predicted means of cultivars when cultivar is fitted as both fixed and random, and when cultivars is fitted as random only. Cultivars Nakitembe, Namwezi and Esenyi show high susceptibility to root knot nematode attack while Salalugazi, and Ndiizi have high resistant to the attack. It is worth noting that cultivar Salalugazi which highly susceptible to weevil attack (discussed in 6.1) is highly resistant to nematode infestation. Cultivar Ndiizi shows high resistant to weevil and nematode infestation in comparison with the other cultivars.

The final models in both analyses seem to retain the same variates. It seems that elevation, rainfall, rainfall prior to the month of sampling,, cultivars, pH, Desuckering and Potassium(K) are the main factors affecting nematode infestation. The coefficients of the final models in tables 6.9 & 6.11 do suggest that, nematode activity increases with increase in elevation, increase of rainfall, increase of rainfall prior to the sampling month, differ in cultivars, decrease of soil pH, increase with Desuckering of banana plants and decrease with increasing of potassium(K) content in the soil.

The sources of soil acidity (pH) include OM (Organic Matter), clay minerals, soluble salts, exchangeable Al^{3+}, and CO_2. Plants require optimum levels of pH for growth (Tisdale ,S. L., Nelson, W. L, (1985)). Soils of high acidity might hinder plant growth while soils of low pH might favor plant growth and hence increase the nematode activity . Plants take the K^+ ion from the soil solution. The concentration of K needed in the soil solution will vary considerably, depending on the type of crop and the amount of growth. The optimum K level in the soil solution is between 10 and 60 ppm, depending on the nature of the crop, soil structure, general fertility level, and moisture supply. Under field conditions, the K concentration of the soil solution varies considerably due to the concentration and dilution processes brought about by evaporation and precipitation, respectively. Diffusion and mass flow of K to plants roots account for the majority of K absorbed. K is essential for root growth hence tissue development (Barber, S. A. 1984). Due to this effect a decrease of K in the soil is likely to weaken roots and hence prone to attack by the root knot nematodes.

The coefficients of the fitted final models seem consistent in sign with few discrepancies in the magnitude. This confirms that the use of cultivars as fixed or random effects makes very little difference. It should further be noted that, the coefficient of rainfall prior to the month of sampling is much bigger in comparison to those of elevation and rainfall during the month of sampling, suggesting a high relationship with the root knot nematode infestation with this variable.

Table 6.9 REML analysis on response variable log(ri+1) with Cultivars as fixed and random effects

No. Of observation=1641

Model 1		Model 2		Model 3		Model 4		Model 5*	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	0.0770	Site	0.0867	Site	0.0901	Site	0.0894	Site	0.0927
Site.farm	0.0756	Site.farm	0.0665	Site.farm	0.0690	Site.farm	0.0688	Site.farm	0.0698
Site.farm.cv	0.0188	Site.farm.cv	0.0173	Site.farm.cv	0.0156	Site.farm.cv	0.0154	Site.farm.cv	0.0153
Units	0.7388	Units	0.7398	Units	0.7401	Units	0.7401	Units	0.7397
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic
Elevation	8.3	Elevation	7.6	Elevation	7.4	Elevation	7.4	Elevation	7.2
Rainfall	4.8	Rainfall	4.7	Rainfall	4.7	Rainfall	4.7	Rainfall	4.7
Prainfall	19.6	Prainfall	19.5	Prainfall	19.5	Prainfall	19.5	Prainfall	19.5
Cultivars	12.2	Cultivars	12.3	Cultivars	12.4	Cultivars	12.4	Cultivars	12.4
Ca	0.0	pH	3.9	pH	3.8	pH	3.9	pH	3.8
pH	4.8	Intercropping	2.4	Intercropping	2.3	Intercropping	2.3	Desuckering	3.2
Intercropping	1.8	Mulching	3.0	Mulching	2.9	Mulching	3.0	K	6.0
Mulching	4.1	De-leafing	0.1	Desuckering	2.5	Desuckering	2.5		
Weeding	0.1	Desuckering	3.3	K	4.0	K	4.0		
De-leafing	1.2	K	4.0	Mg	0.5				
Desuckering	1.7	Mg	1.2						
OM	0.7								
N	0.0								
K	3.4								
Na	0.4								
Sand	0.6								
Clay	0.0								
Silt	0.7								
Mg	1.2								

* Final model:

Y=1.752+0.001235 Elevation + 0.00084 Rainfall + 0.002399 Prainfall + cultivars - 0.1059 pH + 0.1138 Desuckering - 0.07897 K

Important codes : cv - cultivar, Prainfall - Rainfall prior to the month of sampling, OM- Organic Matter

Table 6.10 Table of predicted means for cultivars using log(ri+1) as response variable when cultivars are fitted as both random and fixed effects

Cultivars	Predicted means for log(ri+1)
Nakitembe	1.773
Namwezi	1.768
Ensenyi	1.752
Nfuuka	1.744
Kisubi	1.716
Kisansa	1.698
Likhago	1.664
Kibuzi	1.662
Nassaba	1.638
Mbwazirume	1.606
Nakabululu	1.597
Musakala	1.575
Ndiibwabalangira	1.574
Kayinja	1.568
Nakyetengu	1.550
Ndiizi	1.499
Muvubo	1.465
Salalugazi	1.446

Standard error of differences: 0.1825

Table 6.11 REML analysis on response variable log(ri+1) with Cultivars as Fixed only

No. of observation=1641

Model 1		Model 2		Model 3		Model 4		Model 5*	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	0.0784	Site	0.0871	Site	0.0904	Site	0.0897	Site	0.0928
Site.farm	0.0800	Site.farm	0.0706	Site.farm	0.0725	Site.farm	0.0721	Site.farm	0.0732
Units	0.7521	Units	0.7521	Units	0.7514	Units	0.7513	Units	0.7508
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic
Elevation	8.1	Elevation	7.5	Elevation	7.3	Elevation	7.3	Elevation	7.1
Rainfall	4.7	Rainfall	4.6	Rainfall	4.6	Rainfall	4.6	Rainfall	4.6
Prainfall	19.4	Prainfall	19.3	Prainfall	19.3	Prainfall	19.3	Prainfall	19.3
Cultivars	12.8	Cultivars	12.9	Cultivars	12.9	Cultivars	12.9	Cultivars	12.9
Ca	0.0	pH	3.8	pH	3.8	pH	3.8	pH	3.8
pH	4.7	Intercropping	2.4	Intercropping	2.3	Intercropping	2.3	Desuckering	3.3
Intercropping	1.8	Mulching	3.0	Mulching	2.9	Mulching	2.9	K	5.9
Mulching	4.1	De-leafing	0.0	Desuckering	2.6	Desuckering	2.6		
Weeding	0.2	Desuckering	3.3	K	3.9	K	3.9		
De-leafing	1.1	K	3.9	Mg	0.5				
Desuckering	1.8	Mg	1.2						
OM	0.7								
N	0.0								
K	3.4								
Na	0.4								
Sand	0.6								
Clay	0.1								
Silt	0.6								
Mg	1.2								

* Final model :

Y=1.761+0.001223 Elevation + 0.0008284 Rainfall + 0.002395 Prainfall + Cultivars - 0.1047 pH + 0.1172 Desuckering - 0.07816 K

Important codes: OM - Organic Matter, Prainfall - Rainfall prior to the month of sampling, OM - Organic Matter

6.3 REML analysis of leaf spot data.

6.3.1 Using the variable youngest leaf spotted (YLS) as dependent variable

Chapter 5 discusses leaf spot analysis in two stages, first being the data which only contain those plants which had infestation and secondly the data include the plants which had no leaf spot infestation but the value of no infestation was given a score of 15 instead of 0. This chapter focuses more on the latter data for it seems rather not a good idea to concentrate on trees which had infestation only, which in turn might bias the principal factors affecting this disease. Two approaches have been used to estimate the most important factors influencing leaf spot disease infestation.

First, REML analysis in the mixed model procedure has been used in this chapter to estimate the random and fixed effects using YLS(Youngest leaf spotted with diseases) as the response variable when all the observation with a value of 0 have been given a value of 15, so as to conform with the nature of the infestation as explained in chapter 5. Due to uneven distribution of the response variable (see table 6.17) which might not even be described as skewed in either side of the normal curve and large number of observation with value of 0, which means the sampled plants had no infestation, logistic regression (see section 6.4) was found to be appropriate in analysing the response variable YLS but categorised as 0 for absence of infestation or 1 for presence of disease infestation as discussed in section 6.4. We should bear in mind the limitations of logistic regression as discussed in Chapter 3.

Tables 6.13 and 6.15 shows the results of the mixed models applied on the data and gives the most important factors affecting this infestation in the final models, when using cultivars as fixed and random effects and when using cultivars as fixed effects only. It should be noted that, there is consistent as in the final models in terms of significant factors and the corresponding signs of the coefficients. Model 6* in both tables, shows that elevation, Rainfall prior to the month of sampling (Prainfall) , cultivars, Calcium(Ca), De-leafing and silt seems to the most important factors influencing infestation in one way or another. Tables 6.14 (cultivars fitted as both random

Table 6.13 REML analysis on response youngest leaf spotted with Cultivars as fixed and random effects

No. of observation=2219

Model 1		Model 2		Nodel 3		Model 4		Model 5		Model 6*	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	5.930	Site	5.597	Site	5.479	Site	1.767	Site	5.564	Site	5.811
Site.farm	1.041	Site.farm	0.964	Site.farm	0.953	Site.farm	0.205	Site.farm	0.917	Site.farm	0.888
Site.farm.cv	0.078	Site.farm.cv	0.069	Site.farm.cv	0.069	Site.farm.cv	0.129	Site.farm.cv	0.073	Site.farm.cv	0.079
Units	6.735	Units	6.747	Units	6.749	Units	0.229	Units	6.749	Units	6.752
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic	Fixed term	Wald statistic
Elevation	32.1	Elevation	34.1	Elevation	34.6	Elevation	34.5	Elevation	34.1	Elevation	33.5
Rainfall	2.7	Rainfall	2.7	Rainfall	2.7	Rainfall	2.7	Prainfall	60.8	Prainfall	60.9
Prainfall	58.5	Prainfall	58.4	Prainfall	58.4	Prainfall	58.3	cv	335.4	cv	336.0
Cultivars	327.7	Cultivars	333.3	Cultivars	333.8	Cultivars	335.2	Ca	9.7	Ca	9.8
Ca	8.9	Ca	9.4	Ca	9.5	Ca	9.7	pH	2.2	De-leafing	3.2
pH	2.2	pH	2.3	pH	2.3	pH	2.2	De-leafing	2.3	Silt	5.,9
Intercropp	0.2	De-leafing	2.3	De-leafing	2.3	De-leafing	2.3	Silt	6.9		
Mulching	0.2	OM	0.6	K	1.2	Silt	7.0				
Weeding	0.4	K	2.1	Silt	6.6						
De-leafing	2.7	Clay	0.8								
Desuckering	0.4	Silt	5.8								
OM	1.3										
N	0.3										
K	3.3										
Na	0.0										
Sand	0.6										
Clay	3.1										
Silt	2.2										
Mg	0.1										

* Final model: **Y=9.773+0.01256 Elevation + 0.01083 Prainfall + Cultivars - 0.1277 Calcium + 1.132 De-leafing + 0.05051 Silt**

Important codes: cv - Cultivars, Prainfall- Rainfall prior to the month of sampling, OM - Organic Matter.

Table 6.14: Table of predicted means for cultivars using YLS (Youngest leaf spot) as response variable while cultivars are fitted as both random and fixed effects

Cultivars	Predicted means for YLS
Kayinja	14.242
Kisubi	12.518
Ndiizi	10.857
Ensenyi	9.773
Nakitembe	9.592
Kisansa	9.522
Musakala	9.505
Nfuuka	9.429
Namwezi	9.418
Likhago	9.380
Nakabululu	9.312
Nassaba	9.188
Muvubo	9.170
Mbwazirume	9.039
Ndiibwabalangira	8.959
Kibuzi	8.741
Nakyetengu	8.666
Salalugazi	8.544

Standard error of difference:0.4413

Table 6.15 REML analysis on Youngest leaf spotted as a response variate with cultivars as fixed effects only

No. Of observation=2219

Model 1		Model 2		Model 3		Model 4		Model 5		Model 6 *	
Random term	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**	**Random term**	**Variance Component**
Site	5.942	Site	5.608	Site	5.491	Site	5.573	site	5.578	Site	5.828
Site.farm	1.056	Site.farm	0.976	Site.farm	0.965	Site.farm	0.931	Site.farm	0.931	Site.farm	0.904
Units	6.795	Units	6.799	Units	6.802	Units	6.808	Units	6.805	Units	6.812
Fixed term	Wald Statistic	Fixed term	Wald statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic	Fixed term	Wald Statistic
Elevation	32.1	Elevation	34.0	Elevation	34.6	Elevation	34.4	Elevation	34.4	Elevation	33.4
Rainfall	2.7	Rainfall	2.6	Rainfall	2.6	Rainfall	2.6	Prainfall	60.4	Prainfall	60.4
Prainfall	58.1	Prainfall	58.0	Prainfall	58.0	Prainfall	58.0	Cultivars	346.7	Cultivars	348.2
Cultivars	339.5	Cultivars	343.9	Cultivars	344.4	Cultivars	346.3	Ca	9.7	Ca	9.7
Ca	8.8	Ca	9.3	Ca	9.4	Ca	9.6	pH	2.3	De-leafing	3.1
pH	2.3	pH	2.3	pH	2.3	pH	2.3	De-leafing	2.3	Silt	5.9
Intercropping	0.2	De-leafing	2.2	De-leafing	2.3	De-leafing	2.3	Silt	6.9		
Mulching	0.2	OM	0.6	K	1.2	Silt	6.9				
Weeding	0.4	K	2.2	Silt	6.6						
De-leafing	2.7	Clay	0.8								
Desuckering	0.4	Silt	5.8								
OM	1.3										
N	0.3										
K	3.3										
Na	0.0										
Sand	0.6										
Clay	3.1										
Silt	2.2										
Mg	0.1										

* Final model

Y=9.770 + 0.01257 Elevation + 0.01082 Prainfall + Cultivars - 0.1268 Calcium + 1.121 De-leafing + 0.05034 Silt

Important codes: Prainfall - Rainfall prior to the month of sampling, OM - Organic Matter

Table 6.16 Table of predicted means for cultivars using YLS as response variable and cultivars are fixed effect only

Cultivars	Predicted means for YLS
Kayinja	14.240
Kisubi	12.520
Ndiizi	10.857
Ensenyi	9.770
Nakitembe	9.589
Kisansa	9.524
Musakala	9.502
Nfuuka	9.432
Namwezi	9.409
Likhago	9.372
Nakabululu	9.312
Nassaba	9.201
Muvubo	9.177
Mbwazirume	9.038
Ndiibwabalangira	8.963
Kibuzi	8.736
Nakyetengu	8.673
Salalugazi	8.548

Standard error of difference: 0.4320

6.4 Logistic regression on leaf spot data

As explained in the section 6.3 of this chapter, leaf spot data response variable YLS was further subjected to logistic regression due to

(I) Poor and uneven distribution of this response variable.
(ii) large number of observations having a value of 0 which means no leaf spot infestation, as shown in table 6.17

The logistic regression has limitations in estimating different error structures and in this thesis, it is difficult to estimate the random effects of sites, farms with sites, cultivars within farms while fitting them as both random and fixed effects, unlike the mixed model analysis. Due this statistical problem, we decided to focus separately on data for each sampling visit to minimize error estimations. Table 6.17 gives an overview of the response variable YLS categorised in the form of absence or presence of infestation, it is quite clear that large number of observations have zero responses in overall and in each sampling month. Table 6.18 shows the overall univariate normal distribution of this variable before categorization. The distribution is highly uneven and cannot be even described as skewed.

The logistic regression analysis using the logit link function models the probability that a random sampled plant has leaf spot infection i.e Pr (YLS>0).

Table 6.17 Categorised response variable YLS (0 - absence of infestation; 1 - presence of infestation) by sampling visit

	Overall	January	March	May	July	September	November
YLS=0	826(37.2%)	125(41.9%)	135(32.2%)	124(35.9%)	151(36.7%)	103(27.9%)	170(44.9%)
YLS>1	1393(62.8%)	173(58.1%)	284(67.8%)	221(64.1%)	260(63.3%)	265(72.1%)	208(55.1%)
Total (n)	2219	298	419	345	411	368	378

Table 6.18: Univariate test plot for the variable youngest leaf spotted (YLS)

Univariate Procedure

Variable=YLS

Moments

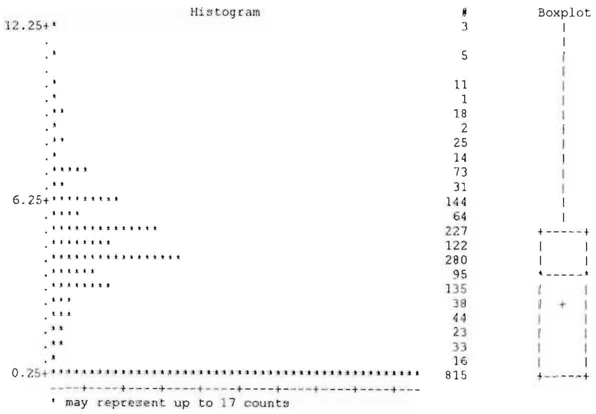
N	2219	Sum Wgts	2219
Mean	2.951385	Sum	6549.123
Std Dev	2.637489	Variance	6.956349
Skewness	0.295698	Kurtosis	-0.8401
USS	34758.17	CSS	15429.18
CV	89.36446	Std Mean	0.05599
T:Mean=0	52.71252	Pr>|T|	0.0001
Num ^= 0	1411	Num > 0	1411
M(Sign)	705.5	Pr>=|M|	0.0001
Sgn Rank	498083	Pr>=|S|	0.0001

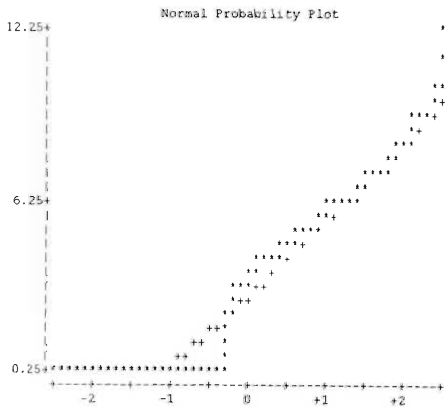
Quantiles(Def=5)

100% Max	12	99%	9.142857
75% Q3	5	95%	7
50% Med	3.5	90%	6
25% Q1	0	10%	0
0% Min	0	5%	0
		1%	0
Range	12		
Q3-Q1	5		
Mode	0		

Extremes

Lowest	Obs	Highest	Obs
0(	2219)	11(	1940)
0(	2218)	11(	1960)
0(	2217)	12(	734)
0(	2216)	12(	1455)
0(	2215)	12(	1541)





6.4.1 Logistic regression using categorized YLS as response variable.

The variable YLS categorised as 0 for absence of infestation and 1 for presence was subjected to logistic regression on data for each sampling visit. Table 6.19 shows the results from logistic regression on data sampled in March, July and November. These tables seem to indicate that elevation, prainfall, cultivars and Ca are the basic principal factors affecting leaf spot infestation. Rainfall during the month of sampling seems not play a major factor in this infestation except possibly in November. This is consistent with the results from the mixed model analysis (see table 6.15). Other factors which seem to contribute to the leaf spot infestation (although to a lesser extent) and are singled out by the logistic analysis are De-leafing and possible pH. These factors are also singled out in the mixed model analysis.

Table 6.20 shows predicted means of the cultivars in logistic analysis with cultivars reacting differently in leaf spot infestation depending on the date of sampling. Salalugazi shows low resistance in March and July sampling but falls to high resistance in the November sampling. The inconsistence of these cultivar differences could be due to the fact that logistic analysis does not account for some of the error structures discussed earlier in this section. Furthermore, by analysis at data for each sampling month separately, the numbers of observations are substantially reduced.

Table 6.19 Logistic regression on data for sampling visits of March, July and November

- Parameter estimates.

	March			July			November		
	estimate	s.e	t(397)	estimate	s.e	t(397)	estimate	s.e	t(397)
Constant	14.4	95.9	0.15	6.	1988.	0.00	4.26	2.39	1.78
elev	-0.00932	.00139	-6.71	-0.00978	.00129	-7.58	-0.00596	0.00122	-4.89
rain	-0.0076	0.0168	-0.45	-0.00765	.00692	-1.11	0.02395	0.00650	3.68
prain	-0.0633	0.0121	-5.23	0.0480	0.0106	4.52	0.02103	0.00994	2.12
cv KAYINJA	0.0	95.9	0.00	2.	1988.	0.00	-3.82	1.89	-2.02
cv KIBUZI	1.4	95.9	0.01	4.	1988.	0.00	-0.29	1.70	-0.17
cv KISANSA	2.7	95.9	0.03	5.	1988.	0.00	-0.31	1.73	-0.18
cv KISUBI	2.2	95.9	0.02	3.	1988.	0.00	-3.53	1.95	-1.81
cv LIKHAGO	1.6	95.9	0.02	5.	1988.	0.00	0.30	1.65	0.18
cv MBWAZIRUME	2.8	95.9	0.03	5.	1988.	0.00	-0.49	1.70	-0.29
cv MUSAKALA	2.0	95.9	0.02	5.	1988.	0.00	-0.31	1.68	-0.18
cv MUVUBO	1.6	95.9	0.02	6.	1988.	0.00	-0.64	1.72	-0.37
cv NAKABULULU	0.9	95.9	0.01	5.	1988.	0.00	-1.31	1.74	-0.75
cv NAKITEMBE	1.8	95.9	0.02	5.	1988.	0.00	-0.38	1.67	-0.23
cv NAKYETENGU	2.0	95.9	0.02	5.	1988.	0.00	-0.12	1.80	-0.07
cv NAMWEZI	2.3	95.9	0.02	13.	2153.	0.01	-0.23	1.95	-0.12
cv NASSABA	7.	100.	0.07	14.	2211.	0.01	4.5	12.9	0.35
cv NDIIBWABALANGIRA	1.9	95.9	0.02	8.	1988.	0.00	0.65	1.75	0.37
cv NDIIZI	0.7	95.9	0.01	5.	1988.	0.00	-2.03	1.84	-1.11
cv NFUUKA	2.8	95.9	0.03	14.	2079.	0.01	0.43	1.71	0.25
cv SALALUGAZI	7.	111.	0.06	14.	2293.	0.01	-1.46	1.88	-0.78
ca	0.1114	0.0379	2.94	0.0505	0.0376	1.34	0.0124	0.0415	0.30

added terms in extended model

ca	0.2833	0.0579	4.89	ca	0.3400	0.0829	4.10	ca	0.0696	0.0534	1.30
dl1	-2.235	0.551	-4.06	dl1	-1.927	0.701	-2.75	dl1	-3.059	0.773	-3.96
silt	-0.0792	0.0338	-2.34	ph	-2.498	0.539	-4.63	ph	0.554	0.269	2.06
pH	-0.996	0.286	-3.49					silt	-0.0657	0.0329	-2.00
clay	-0.0153	0.0173	-0.88					clay	0.0452	0.0177	2.56

Table 6.20 Logistic regression on data for sampling visits of March, July and November

- Cultivar means.

	Month 3		Month 7		Month 11	
	Prediction	S.e.	Prediction	S.e.	Prediction	S.e
ENSENYI	0.3545	21.9477	0.0134	26.3596	0.7302	0.3246
KAYINJA	0.3626	0.1018	0.1104	0.0536	0.0560	0.0405
KIBUZI	0.6888	0.0920	0.4543	0.1033	0.6693	0.1127
KISANSA	0.8900	0.0741	0.7441	0.1401	0.6641	0.1208
KISUBI	0.8320	0.1538	0.1610	0.1067	0.0733	0.0636
LIKHAGO	0.7258	0.0641	0.7347	0.0738	0.7848	0.0579
MBWAZIRUME	0.8978	0.0691	0.7440	0.1046	0.6232	0.1193
MUSAKALA	0.8095	0.0795	0.5938	0.0916	0.6649	0.0898
MUVUBO	0.7226	0.0954	0.7713	0.0958	0.5877	0.1309
NAKABULULU	0.5624	0.0923	0.6069	0.0924	0.4216	0.1548
NAKITEMBE	0.7767	0.0676	0.6375	0.0802	0.6493	0.0833
NAKYETENGU	0.8053	0.1379	0.6461	0.1476	0.7065	0.1614
NAMWEZI	0.8516	0.1780	0.9999	0.1158	0.6818	0.2253
NASSABA	0.9984	0.0473	0.9999	0.0521	0.9960	0.0514
NDIIBWABALANGIRA	0.7895	0.1018	0.9639	0.0416	0.8381	0.0928
NDIIZI	0.5345	0.0993	0.6365	0.0894	0.2621	0.1472
NFUUKA	0.8998	0.0738	0.9999	0.0457	0.8065	0.0790
SALALUGAZI	0.9976	0.1343	0.9999	0.0763	0.3852	0.2113

CHAPTER 7

CONCLUSIONS

7.1 Statistical Models

Statistical models for data are mathematical descriptions of how data conceivably can be produced. Mixed model analysis has been described by Singer, J.D.,(1997); Latour, D., Latour, K. & Wolfinger, R.D (1994) as an increasingly popular method in fitting multilevel models, hierarchical models and individual growth models in experimental designs, sample survey, or an observational study.

In this study, multilevel and hierarchical models have been fitted to explain various infestation effects involving banana pests and diseases. Due the error structures involved in the sample survey, we emphasize the use of mixed models applied in this data, as the most appropriate method of estimating and isolating various factors affecting infestation. We believe the methods used here could be taken as a bench mark in analysing future surveys for the purpose of verification and comparing of the results with the final models describing each of the three infestation types (Weevils, Nematodes and Leaf spot) . Naturally the model parameter estimates obtained are subject to sampling errors and sampling judgmental errors from the enumerators. Nevertheless, mixed model analysis does minimize the sampling errors by taking into account variance components from the multilevel strata

7.1.1 Weevil infestation models

The models explaining weevil infestation attempt to identify the major factors affecting the infestation. Mixed model analyses have singled out Elevation, Rainfall, Prainfall and Cultivars, to be the major factors and probably Na and N as well. We point the absence of farm management practices in our final models describing this infestation and wish to state that, further investigation

Table 6.20 Logistic regression on data for sampling visits of March, July and November

- Cultivar means.

	Month 3		Month 7		Month 11	
	Prediction	S.e.	Prediction	S.e.	Prediction	S.e
ENSENYI	0.3545	21.9477	0.0134	26.3596	0.7302	0.3246
KAYINJA	0.3626	0.1018	0.1104	0.0536	0.0560	0.0405
KIBUZI	0.6888	0.0920	0.4543	0.1033	0.6693	0.1127
KISANSA	0.8900	0.0741	0.7441	0.1401	0.6641	0.1208
KISUBI	0.8320	0.1538	0.1610	0.1067	0.0733	0.0636
LIKHAGO	0.7258	0.0641	0.7347	0.0738	0.7848	0.0579
MBWAZIRUME	0.8978	0.0691	0.7440	0.1046	0.6232	0.1193
MUSAKALA	0.8095	0.0795	0.5938	0.0916	0.6649	0.0898
MUVUBO	0.7226	0.0954	0.7713	0.0958	0.5877	0.1309
NAKABULULU	0.5624	0.0923	0.6069	0.0924	0.4216	0.1548
NAKITEMBE	0.7767	0.0676	0.6375	0.0802	0.6493	0.0833
NAKYETENGU	0.8053	0.1379	0.6461	0.1476	0.7065	0.1614
NAMWEZI	0.8516	0.1780	0.9999	0.1158	0.6818	0.2253
NASSABA	0.9984	0.0473	0.9999	0.0521	0.9960	0.0514
NDIIBWABALANGIRA	0.7895	0.1018	0.9639	0.0416	0.8381	0.0928
NDIIZI	0.5345	0.0993	0.6365	0.0894	0.2621	0.1472
NFUUKA	0.8998	0.0738	0.9999	0.0457	0.8065	0.0790
SALALUGAZI	0.9976	0.1343	0.9999	0.0763	0.3852	0.2113

CHAPTER 7

CONCLUSIONS

7.1 Statistical Models

Statistical models for data are mathematical descriptions of how data conceivably can be produced. Mixed model analysis has been described by Singer, J.D.,(1997); Latour, D., Latour, K. & Wolfinger, R.D (1994) as an increasingly popular method in fitting multilevel models, hierarchical models and individual growth models in experimental designs, sample survey, or an observational study.

In this study, multilevel and hierarchical models have been fitted to explain various infestation effects involving banana pests and diseases. Due the error structures involved in the sample survey, we emphasize the use of mixed models applied in this data, as the most appropriate method of estimating and isolating various factors affecting infestation. We believe the methods used here could be taken as a bench mark in analysing future surveys for the purpose of verification and comparing of the results with the final models describing each of the three infestation types (Weevils, Nematodes and Leaf spot) . Naturally the model parameter estimates obtained are subject to sampling errors and sampling judgmental errors from the enumerators. Nevertheless, mixed model analysis does minimize the sampling errors by taking into account variance components from the multilevel strata

7.1.1 Weevil infestation models

The models explaining weevil infestation attempt to identify the major factors affecting the infestation. Mixed model analyses have singled out Elevation, Rainfall, Prainfall and Cultivars, to be the major factors and probably Na and N as well. We point the absence of farm management practices in our final models describing this infestation and wish to state that, further investigation need to be carried out either by a field experimental trial or a well design sample survey. Rukazambuga (1997) described mulching as a major factor affecting yield loss. Farms which had high mulching and infested with weevils, had high overall yield loss even though the yield production was high compared to the plots without mulching. The soil factors need further to be studied for soil texture and nutrients change with season, a factor which was not taken care of during the survey, for data was only collected once and our assumptions are that, these nutrients remain almost the same through out the entire period of the survey.

Rainfall prior to the sampling period seems to influence weevil infestation more than the rainfall during the sampling period. Further studies need to be initiated to verify and compare these results for no literature so far available on the effect of this factor. Table 7.1 gives a summary of factors influencing weevil infestation from various statistical methods used in this thesis while table 7.2 shows cultivar susceptibility from different analyses carried out.

Table 7.1 Summary of factors affecting weevil infestation by various analysis

	MANOVA		REML			
	Untransformed	Log transformation	PC1	(sign)*	Log (xt+1)	(sign)*
Elevation	**	**	**	-	**	-
Rainfall	**	**	**	-	**	-
Prainfall	NS	NS	**	-	**	-
Cultivars	**	**	**		**	
K	**	**	NS		NS	
Ca	**	NS	NS		NS	
De-leafing	NS	*	NS		NS	
Na	NS	NS	*	-	NS	
N	NS	NS	**	-	*	-

* A positive sign implies that increasing this factor will increase the weevil infestation score and a negative sign conversely.

Table 7.2 Summary of cultivar difference in weevil susceptibility

	Correspondence Row score	REML PC 1 (Predicted means)	REML log (xt+1) - Predicted means
Kisubi	-1.459	0.899	0.020
Kayinja	-0.994	1.493	0.349
Ndiizi	-0.987	1.734	0.450
Ensenyi	-1.119	3.113	0.856
Kibuzi	-0.070	3.855	1.141
Mbwazirume	-0.072	3.674	1.103
Nakabululu	0.038	3.708	1.136
Musakala	0.056	3.577	1.104
Ndiibwabalangira	0.135	3.472	1.045
Nfuuka	0.179	3.643	1.077
Nakyetengu	0.122	3.023	0.897
Muvubo	0.165	3.463	1.025
Nakitembe	0.248	3.705	1.131
Likhago	0.303	3.713	1.113
Namwezi	0.256	3.744	1.119
Kisansa	0.487	3.746	1.136
Nassaba	0.487	4.373	1.366
Salalugazi	0.658	4.529	1.435

7.1.2 Nematode infestation models

Infestation due to this pest on bananas has been explained by various models. Mixed models seem to suggest a wide range of factors influencing the infestation in different ways. This includes ecological factors and farm management practices. Again Prainfall (Rainfall prior to the month of sampling) seems to come in strongly as a major factor affecting nematode infestation in addition to the factors (Elevation, Rainfall, Cultivars, pH, Desuckering and K). Since no literature available focusing on the accumulated rainfall effects over a period of time further studies in a more controlled experiment or sample survey is needed to verify these results. Soil nutrients pH and K in the final model suggest that, soil nutrients play a major role in infestation and a detailed study covering this aspect would be valuable.

Tables 7.3 and 7.4 gives a summary of factors influencing nematode infestation and cultivar susceptibility as derived from different analysis. These results seem to imply that there are small differences between cultivars in their nematode susceptibility.

Table 7.3 Summary of factors influencing nematode infestations by various analysis

	REML (when cultivars are fitted as both random and fixed) on log (RI+1)	Sign*	REML (when cultivars are fitted as and fixed only) on log (RI+1)	Sign*
Elevation	**	+	**	+
Rainfall	**	+	**	+
Prainfall	**	+	**	+
Cultivar	**		**	
pH	*	-	*	-
Desuckering	*	+	*	+
K	**	-	**	-

* A positive sign implies that increasing this factor will increase the nematode infestation score and a negative sign conversely.

Table 7.4 Summary of cultivar difference in nematode susceptibility

	Correspondence analysis (Row scores)	REML (Cultivar as fixed & random) - Predicted means	REML (Cultivar as fixed & random) - Predicted means
Mbwazirume	-0.285	1.606	1.614
Musakala	-0.103	1.575	1.579
Ndiizi	-0.137	1.499	1.496
Nakabululu	-0.278	1.597	1.601
Muvubo	-0.143	1.465	1.463
Kibuzi	-0.139	1.662	1.665
Likhago	0.079	1.664	1.666
Kisansa	0.105	1.698	1.696
Ndiimbwabalngira	0.028	1.574	1.577
Ensenyi	0.028	1.752	1.761
Nakitembe	0.068	1.773	1.774
Salalugazi	0.173	1.446	1.456
Nassaba	0.184	1.638	1.646
Nfuuka	0.181	1.744	1.747
Kayinja	0.194	1.568	1.555
Nakyetengu	0.237	1.499	1.553
Kisubi	0.098	1.716	1.719
Namwezi	0.378	1.768	1.762

7.1.3 Leaf spot infestation models

Several statistical methods have been needed to explain the factors influencing leaf spot infestation. The reason being the nature of the data collected where a large number of the observations turned out to be 0 (no infestation). Mixed model analyses and logistic analyses have been used. The most important factors affecting this disease resulting from our models are Elevation, Prainfall, Cultivars and Ca . These results are quite consistent on both from mixed model analysis and logistic analysis although the logistic analyses is less valuable.

Tables 7.5 and 7.6 gives an summary of statistical analysis used to quantify this infestation and cultivar resistance to this disease as derived from different statistical analysis.

Table 7.5 Summary of factors influencing leaf spot infestation by various analyses

	REML (Cultivars as fixed & Random)	Sign*	REML (Cultivars as fixed only)	Sign*	Logistic regression analyses	Sign*
Elevation	**	-	**	-	**	-
Rainfall	NS		NS		NS	
Prainfall	**	-	**	-	NS	
Cultivars	**		**		**	
Ca	**	+	**	+	*	+
De-leafing	*	-	*	-	*	-
Silt	**	-	**	-	NS	
pH	NS		NS		*	-

* A positive sign implies that increasing this factor will increase the leaf spot infestation score and a negative sign conversely.

Table 7.6 Summary of cultivar difference in leaf spot resistance

	Correspondence analysis (Row scores) *	REML (Predicted means) *	Logistic regression (Predicted means) *		
			March	July	Nov
Salalugazi	0.781	8.544	0.997	0.999	0.385
Namwezi	0.680	9.418	0.851	0.999	0.682
Nassaba	0.441	9.188	0.998	0.999	0.996
Ndiibwabalangira	0.432	8.959	0.789	0.963	0.838
Nfuuka	0.398	9.429	0.899	0.999	0.806
Kisansa	0.366	9.522	0.890	0.744	0.664
Nakyetengu	0.362	8.666	0.805	0.646	0.706
Ensenyi	0.215	9.773	0.354	0.013	0.730
Muvubo	0.202	9.170	0.722	0.771	0.587
Musakala	0.050	9.505	0.809	0.593	0.665
Mbwazirume	-0.044	9.039	0.897	0.744	0.623
Kibuzi	-0.076	8.741	0.688	0.454	0.669
Nakitembe	-0.123	9.592	0.776	0.637	0.649
Likhago	-0.129	9.380	0.725	0.734	0.784
Nakabululu	-0.245	9.312	0.562	0.606	0.421
Ndiizi	-0.443	10.857	0.534	0.636	0.262
Kisubi	-0.708	12.518	0.832	0.161	0.664
Kayinja	-0.873	14.242	0.363	0.110	0.056

* Note that an increase in leaf spot susceptibility is associated with a decrease in correspondence analysis row scores, a decrease in the REML means and an increase in logistic regression means.

REFERENCES

Agresti, A. (1984). Analysis of Ordinal Categorical Data, New York: John Wiley & Sons, Inc.

Anderg, M.R. (1973). Cluster analysis for applications, New York: Academic Press, Inc.

Anderson, T.W. (1958). An introduction to Multivariate Statistical Analysis, New York: John Wiley & Sons, Inc.

Barber, S. A. (1984). Soil Nutrient Bioavailability: A Mechanistic Approach. Joney Wiley & Sons, New York.

Beaton, J.D., Fox, R. L., and Jones, M. B. (1985). Production, marketing, and use of sulphur products, pp. 411-454. In O. O. Englestad (Ed.) , Fertilizer Technology and use. Soil Science Society of America, Madison, Wisc.

Benzécri, J.P. (1973). L´Analyse des Données: T. 2, L´Analyse des Correspondances, Paris: Dunod .

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975). Discrete Multivariate Analysis: Theory and Practice, Cambridge, MA: The MIT press.

Breslow, N.E & Clayton, D.G (1983). Approximation inference in generalized linear mixed models. JASA, 88, 421, 9-25

Burt. C. (1950). The Factorial Analysis of Qualitative Data. British Journal of Psychology, 3, 166-185.

Byran F.J Manly (1994), Multivariate Statistical Methods, 2nd Edition, New Delhi: Thomson press

Cole, J.W. L. and Grizzle, J.E. (1966). Applications of Multivariate Analysis of Variance to Repeated Measures Experiments. Biometrics, 22, 810-828.

Cooley, W.W. and Lohnes, P.R. (1971). Multivariate Data Analysis, New York: John Wiley & Sons, Inc.

Cooley, W.W and Lohnes, P.R(1971), Multivariate Data Analysis, New York: John Wiley & Sons, Inc.

Cooper, M.C. and Milligan, G,W. (1984). The effect of Error on Determining the Number of Clusters. College of Administrative Science Working Paper Series 84-2, Columbus, OH: The Ohio State University.

Daultry, S. (1976). Principal Component Analysis. Concepts and Techniques of Modern Geography, &, Geo. University of East Anglia, UK.

Everitt, B.S. (1980). Cluster analysis, 2nd Edition, London: Heineman Educational Books Ltd.

Everitt, B.S. (1995), "The analysis of Repeated Measures: A Practical Review with Examples," The *Statistician*, 44, 113-135

Fisher, R.A. (1936). The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, 7, 179-188.

Fisher, R.A. (1940). The Precision of Discriminant Functions. Annals of Eugenics, 10, 422-429.

Follet, R. H., Murphy, L. S., and Donahue, R. L. (1981). Fertilizers and soil amendment. *Printice-Hall, Inc., Englewood Cliffs, N.J*

Forthofer, R.N. and Koch, G.G. (1973). An Analysis of Compounded Functions of Categorical Data. Biometrics, 29, 143-157.

Furry B. (1951). "A first course in multivariate statistics". Springer-verlag New York, Inc.

GENSTAT 5 (1989) Release 3 Reference Manual. Oxford University Press Inc., New York

Gilmour, A.R., Thompson, R. And Cullis, B. (1995). AIREML, an effecient algorith for varaince parameter estimation in linear mixed models.Biometrics, 51, in Press

Gnanadesikan, R. (1977). Methods for Statistical Data Analysis of Multivariate Observations, new York: John Wiley & Sons, Inc

Gold, C.S., Ogenga-latigo,M.W., Tushemereirwe,W.K., Kashaija, I. AND Nankinga,C.(1993): Farmer perceptions of banana pest constraints in Uganda: Farmer perceptions of banana pest contraints in Uganda: Results from a rapid rural appraisal. In: *Proceedings of a Research Coordination Meeting for Biological and Integrated Control of Highland Banana and Plantain Pests and Diseases*. Cotonou, 12-14 November 1991. pp.3-24.(eds Gold, C.S. and Gemmil, B.). IITA, Cotonou, Benin.

Gold, C.S., Ogenga-Latigo, M.W., Tushemereirwe, W. K., Kashaija, I.N., and Nankinga, C. (1993). Farmer perceptuons of banana pest constraints in Uganda: Results from a rapid rural appraisal. In *Proceedings of a Research Coordination Meeting for Biological and Integrated Control of Highland Banana Pests and Diseases in Africa.* Cotonou, 12-14 November 1993. C.S. Gold and B. Gemmill (edit.), pp 3-24. IITA, Ibadan, Nigeria

Gold C.S. , Okech S.H., Karamura E.B. & Ssendege R. (1997). Banana weevil population densities and related damage in Ntungamo and Mbarara districts, Uganda. African Crop Science Conference Proceedings. Vol 3. Pp. 1207-1219

Gold, C.S., Speijer, P.R (1994). Assessment of banana weevils in East African Highland banana systems and strategies for control. Pages 170-190. In: R.V. Valmayor, R.G. Davide, J.M. Stanton, N.L. Treverrow & V.N Rao (Eds.), Banana Nematodes and Weevil Borers in Asia abd the Pacific. Proceedings of a Conference-Workshop on Nematodes and Weevil Borers Affecting Bananas in the Asia and Pacific, Serdang, Selangor, Malayasia, 18-22 April 1994.

Gower, J.C. (1967). A Comparison of Some Methods of Cluster Analysis. Biometrics, 23, 623-637.

Greenacre, M.J. (1984). Theory and applications of correspondence analysis. Academic Press, London.

Guthrie, D. (1981). Analysis of Dichotomous Variables in Repeated Measures Experiments. Phsychological Bulletin, 90, 189-195

Guttman, L. (1941). The Quantification of a Class of Attributes: A Theory and Method of Scale Construction. In the Prediction of Personal Adjustment, eds. P. Horst, et al., New York: Social Science Research Council.

Hayashi, C. (1950). On the Quantification of Qualitative Data from the Mathematico-Statistical Point of View. Annals of the Institute of Statistical Mathematica, 2, (No. 1), 35-47.

Henderson, C.R. (1984). Applications of Linear Models in Animal Breeding, University of Guelph.

Hirshfied, H.O. (1935). A Connection Between Correlation and Contingency. Cambridge Philosophical Society Proceedings, 31, 520-524.

Hochberg, Y. (1974). Some Conservative Generalizations of the T-Method in Simultaneous Inference. Journal of Multivariate Analysis, 4, 224-234.

Horst, P. (1935). Measuring Complex Attitudes. Journal of Social Phychology, 6, 369-374.

Hotelling, H. (1933). Analysis of complex of statistical variables into principal components. Journal of Educational Physchology 24, 417-41, 498-520.

Hotelling, H. (1936), "Relations between two sets of variables," Biometrika, 28, 321-377

Huynh, H. And Feldt, L.S. (1970). Conditions under Whicn Means Square Ratios in Repeated Measurements Designs Have Exact F-Distributions. Journal of the American Statistical Association, 65, 1582-1589.

INIBAP. (1986). A preliminary study of the needs for banana research in Eastern Africa. INIBAP doc. EA-001e.

Jagtap, S.S. (1993). Site selection using GIS for effective biological and integrated control of highland banana pests. In *Proceedings of a Research Coordination Meeting for Biological and Integrated Control of Highland Banana Pests and Diseases in Africa.* Cotonou, 12-14 November. C.S. Gold and B. Gemmill (edit), pp 25-36. IITA. Ibadan, Nigeria.

Johnson, A. J., Dean, W.W. (1982). Applied Multivariate Statistical Analyses, pp. 252-258 . Englewood Cliffs, Inc.

Karamura, E.B. 1992. Banana/plantain production constraints as a basis for selecting research priorities. In: Proceedings of the Regional Advisory Committee Meeting. Regional Network for Eastern Africa . Kampala, Uganda 23-25 September 1991. INIBAP, Montpelier.

Kashaija, I. and Speijer, P.R., Gold, C.S., and Gowen, S.R.1994. Occurence, distribution and abundance of plant parasitic nematodes on bananas in Uganda: preliminary results of a diagnostic survey. *African crop science Journal* 2 (1): 99-104

Krishnaiah, P.R and Armitage, J. V. (1966). Tables for Multivariate t Distribution. Sankya, series B, 31-56

Kshirgsagar, A.M. (1972). Multivariate Analysis, New York: Marcel Dekker, Inc.

Latour, D., Latour, K. & Wolfinger, R.D (1994). Getting started with PROC MIXED. Cary, NC:SAS Institute, Inc.

LaTour, S.A. and Miniard, P.W. (1983). The Misuse of Repeated Measures Analysis in Marketing Research. Journal of Marketing Research, XX, 45-47.

Lawley, D.N(1959), "Tests of Significance in Canonical Analysis, "Biometrika, 46, 59-66.

Littel, R.C., Miliken, G.A., Stroup, W.W., and Wolfinger, Russel D., (1996). SAS System for Mixed Models, Cary, NC: SAS Institute Inc., 1996.633 pp.

Littel, R.C., Milliken, G.A., Stroup, W.W., & Wolfinger, R.D. (1996). SAS System for Mixed Models. Cary, NC: SAS Institute, Inc.

Mardia, K.V., Kent, J.T., and Bibby, J.M. (1979). Multivariate analysis, London: Academic Press.

McCullagh, P. & Nelder, J.A(1989). Genelarized Linear Models, Second Edition. Chapman & Hall, London.

Mead, R., Curnow, R.N., and Hasted, A.M (1993). Statistical Methods in Agriculture and Experimental Biology. Second Edition. Chapman & Hall.

Mitchell, G.(1978). *The estimation of banana borer population and resistance levels*. Windward Banana Research and Development Technical Bulletin 2. pp.34.

Mitchell, G.(1980). *Banana Entomology in the Windward Islands*.Final Report 1974-1978. pp.216 Centre for overseas pest research, London.

Morrison, D.F. (1976), Multivariate Statistical Methods, 2nd Edition, New York: McGraw-Hill Book Co.

Ngoya J.N., Clarke G.P.Y, Gold C.S., Speijer P., (1998) "Mixed models applied on Survey data set to explain weevil, nematode and leaf spot infestation". *In the proceeding of the International Biometrics Conference.*

Pearson, K. (1901). On lines and Planes of Closest Fit to Systems of Points in Space.Philosophical Magazine, 6(2), 559-572.

Pinochet, J. 1988, A method of screening bananas and plantains to lesions forming nematodes. 62-65 in Nematodes and Borer Weevil in Bananas: Present Status of Research and Outlook. INIBAP, Montpellier, France.

Power, J. F., and Papendick, R. I. (1985). Organic sources of nitrgon, pp. 503-520. In O.P. Englelstad (Ed.). Fertilizer Technology and Use. Soil Science Society of America. Madison, Wisc.

Price, N.(1995). Banana morphology - part I: Roots and rhizomes.In: *Bananas and plantains*. pp. 179 - 189.(ed. Gowen, S.). Chapman and Hall, London.

Rao, C.R(1972), "Estimation of Variance and Covariance Components in linear models," JASA, 67, 112-115.

Rao, C.R(1964), "The use and interpretation of Principal Component Analysis in applied research," Sankya A, 26, 329-358

Richard, A. J, and Dean W.W. (1982). Applied Multivariate Statistical Analysis. Prentice-Hall International, Inc., London

Richardson, M. And Kuder, G.F. (1933). "Making a Rating Scale that Measures". Personnel Journal, 12, 36-40.

Robinson J.C. (1993). Handbook of Banana growing in South Africa. Institute of Tropical and Subtropical Crops.

Rubin, D.B. (1976). Inference and Missing Data. Biometrika, 63, 112-115.

SAS Institute Inc., (1989). SAS/STAT "*Users Guide, Version 6, Fourth Edition, Volume* 1, Cary, NC:SAS Institute Inc.

SAS Institute Inc. (1990). SAS Procedures Guide, Version 6, Third Edition, Cary, NC: SAS Institute Inc.

SAS Institute Inc. (1990). SAS Language Guide: Reference Manual, Version 6, First Edition, Cary, NC: SAS Institute Inc.

Schall , R. (1991). Estimation in Generalized linear models with random effects. Biometrika, 78,719-727.

Searl, S.R (1971). Linear Models, New York: John Wiley & Sons, Inc.

Searl, S.R. , Casella, G., and McCulloch, C.E (1992). Variance Components. New York: John Wiley & Sons, Inc.

Sebasigari, K. and Stover,R.H. 1988, Banana disease and pests in East Africa: A report of a survey made in November, 1987, INIBAP, 15 PP.

Self, S.G. and Liang, K.Y (1987). Asymptotic Properties of Maximum Likelihood Estimaters and Likelihood Ratio Tests Under Nonstandard Conditions. Journal of the American Statistical Association, 82, 605-610.

Sikora, R.A. Bafokuzara, N.D. Mbwana, A.S.S., Oloo, G.W., Uronu, B. and Seshu Reddy, K.V. (1989). Interrelationships between banana weevil, root lesion nematode and agronomic practices and their importance for banana decline in Tanzania. F.A.O. Plant Protection Bulletin. 37:151-157

Simmonds, N.W.(1966). *Bananas*. pp.512. Longmans, London

Singer J.D. (1997). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Harvard University. Cambridge

Snedecor, G.W. and Cochran, W.G. (1980). Statistical Methods, Ames: The Iowa State University Press.

Speijer, P.R., Gold, C.S., Karamura, E.B and Kashaija, I.N. (1994b). Assessement of banana nematodes in East African highland banana systems. In: *Nematode and Weevil Borer Conference Proceedings.* Kuala Lumpur, 19-23 April 1994.

Speijer, P.R. , Gold, C.S., Karamura, E.B. and Kashaija. I.N. (1994a). Banana weevil and nematode constraints in Uganda: Preliminary results from a diagonistic survey. In E. Adipala (Edit). *Proceedings of the First Annual Crop Science Conference, 14-18 June 1993,* Kampala Uganda.

Speijer, P.S., and de Waele D. 1997, Screening of Musa Germplasm for resistance and tolerance to nematodes. INIBAP, Technical Guidence, 42 pp.

Stover, R.H. and Simmonds, N.W. (1987). *Bananas*. pp.468. Longman, London.

Stover, R.H. (1972). Banana plantain and abaca diseases. Kew: Commonwealth Mycological Institute. London. 316 pp.

Tisdale S. L, Nelson, W. L., Beaton, J. D, Havlin, J. L. (1985). Soil fertility and fertilizers., pp. 111-175. Macmillan publishing company. New York.

Tothill, H. (1940). *Bananas*. pp.120. Longman, London.

Tushemereirwe, K.W. 1996. *Factors influencing the expression of leaf spot diseases of highland bananas in Uganda*. Ph.D. thesis. Reading University.

Uganda Ministry of Agriculture. 1992. *Report on Ugandan National Census of Agriculture and Livestock* (1990-1991). 82 pp + 40 tables. Entebbe, Uganda.

Winer, B.J. (1971). Statistical Principles in Experimental Design, Second Edition. New York: McGraw-Hill, Inc

DIAGNOSTIC SURVEY
AGRONOMIC CHECKLIST

SITE FARM

DATE INITIALS

A. INTERCROPPING ___________

USE 50:50 INTERCROP RATIO AS BASIS FOR COMPARISON
INQUIRE ABOUT INTERCROPPING DURING OTHER SEASONS

0 YEAR-ROUND MONOCULTURE
1 SPARSE OR PARTIAL (MODERATE) ANNUAL INTERCROPS
2 SPARSE OR PARTIAL (MODERATE) PERENNIAL INTERCROPS
3 MODERATE ANNUAL INTERCROPS THROUGHOUT FIELD
4 MODERATE PERENNIAL INTERCROPS THROUGHOUT FIELD
5 < 33% OF FIELD INTENSIVELY INTERCROPPED WITH ANNUALS
6 < 33% OF FIELD INTENSIVELY INTERCROPPED WITH PERENNIALS
7 33 - 67 % OF FIELD INTENSIVELY INTERCROPPED WITH ANNUALS
8 33 - 67 % OF FIELD INTENSIVELY INTERCROPPED WITH PERENNIALS
9 > 67 % OF FIELD INTENSIVELY INTERCROPPED WITH ANNUALS
10 > 67 % OF FIELD INTENSIVELY INTERCROPPED WITH PERENNIALS

INTENSIVE INTERCROPS

MODERATE INTERCROPS

CYCLES OF ANNUAL INTERCROPS

DIAGNOSTIC SURVEY
AGRONOMIC CHECKLIST

SITE FARM

DATE INITIALS

B. MULCHING ____________

0	NO MULCH	
1	LIGHT BANANA TRASH MULCH	(< 25% FIELD COVERED)
2	LIGHT BEAN RESIDUES	
3	LIGHT MAIZE STALK RESIDUES	
4	MODERATE BANANA TRASH MULCH	(25-75% FIELD COVERED)
5	MODERATE BEAN RESIDUES	
6	MODERATE MAIZE STALK RESIDUES	
7	MODERATE IMPORTED GRASS MULCH	
8	HEAVY BANANA TRASH MULCH	(>75% FIELD COVERED < 1")
9	INTENSIVE BANANA TRASH MULCH	(>75% FIELD COVERED > 1")
10	CUT AND SPREAD PSEUDOSTEMS	
11	HEAVY IMPORTED GRASS MULCH	
12	OTHER L/M/H	TYPE MULCH

AVERAGE DEPTH OF MULCH:

MULCH TO BASE OF PLANTS YES/NO

C. DELEAFING:

Frequency:

Rainy Season	Dry Season	Per Year	Piecemeal
----------	----------	--------	----------

Selection:

Reason:

D. DESUCKERING

Frequency:

Rainy Season	Dry Season	Per Year	Piecemeal
----------	----------	--------	----------

Selection:

Reason:

(Draft Study Material Not Fof Citation)

WEEVIL ASSESSMENT: INTENSIVE STUDY SITES

SITE FARM CODE DATE INITIALS

	PLANT MAT	PLANT PLT	CULTIVAR	X Sections UI	UO	LI	LO	FUS	GIR	AGE
1.										
2.										
3.										
4.										
5.										
6.										
7.										
8.										
9.										
10.										
11.										
12.										
13.										
14.										
15.										
16.										
17.										
18.										
19.										
20.										
21.										
22.										

U: UPPER; L: LOWER; I: INNER; O: OUTER

BANANA WEEVIL PCI SCORES

SITE
DATE

FARM CODE
ENUMERATORS

COORDINATES
ELEVATION

		H/T		PCI SCORES SECTION										X SECTION UPPER		LOWER	
	CULTIV.	S/D	GIR	1	2	3	4	5	6	7	8	9	10	I	O	I	O
1																	
2																	
3																	
4																	
5																	
6																	
7																	
8																	
9																	
10																	
11																	
12																	
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	

H: HARVESTED; T: TOPPLED; S: SNAPPED; D: DEAD
GIR: GIRTH AT GROUND LEVEL