

**UNIVERSITY OF NAIROBI**



**NAIROBI UNIVERSITY  
KABETE LIBRARY**

**QUANTIFYING AND MAPPING SUGARCANE BIOMASS USING REMOTE  
SENSING**

**MASTERS THESIS**

**By**

**ORODI ODHIAMBO JOHANNES DOVENS**

**Registration No. F56/8506/00**

A thesis submitted to the Department of Environmental and Biosystems Engineering  
in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE**  
**IN ENVIRONMENTAL AND BIOSYSTEMS ENGINEERING**

**JANUARY 2008**

University of NAIROBI Library



0524693 9

**NAIROBI UNIVERSITY  
KABETE LIBRARY**

## DECLARATION

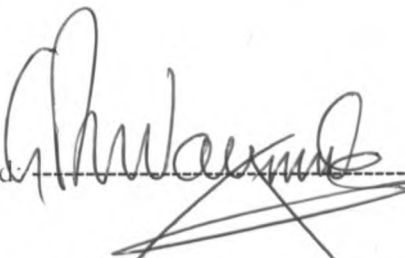
I declare that this thesis is my original work and has not been presented in any university. All sources of information and literature cited have been referenced and acknowledged accordingly.

Signed  Date 21/01/2008

**Orodhi Odhiambo Johannes Dovens**

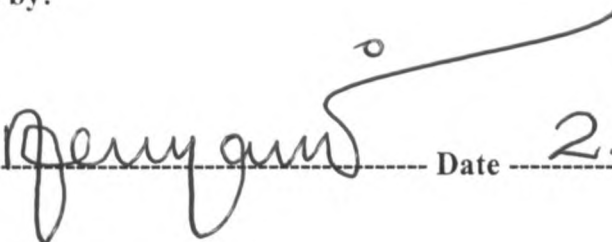
This thesis has been submitted for examination with our approval as the University Supervisors.

Approved by:

Signed  Date 22nd January 2008

**Mr. Gordon Okumu Wayumba**  
First Supervisor

Approved by:

Signed  Date 25.01.08

**Mr. Albert K. Inima**  
Second Supervisor

## QUOTES

**“For these seven rejoice to see... for they are the eyes of the LORD, which scan to and fro throughout the whole earth”**

**Zechariah 4:10 (KJV).**

*The true genesis of remote sensing as carved in the holy writs*

**The most beautiful thing we can experience is the mysterious. It is the source of all true art and science.**

*Albert Einstein*

**Herein is the mystery of godliness, being so distant yet so near, so removed yet omnipresent, the thrust and throb of remote sensing technology.**

Kwe

## DEDICATION

To Jesus Christ my LORD, the alpha and omega of all creation, who controls and manages all things from so far yet, so near;

To Wisdom Joy Doreen Kwe, my daughter and the progeny of my being who is ever near and a constant loving reminder of my parenthood;

To Emily Adhiambo, for the joy of her youthful love that has brought brightness to our hearts;

To Nyar Alego and Jak'Orende, my parents, for their desire that I excel in all my endeavors;

And to those who aspire to be in touch without manifesting in person.

To you who cherish the manifestation of the mysterious, I dedicate this work.

## ACKNOWLEDGMENTS

I would like to express my gratitude to the following people and instructors for their contribution towards the completion of this thesis.

Prof. Francis W. Aduol, the Principal, College of Architecture and Engineering, University of Nairobi, deserves my sincere thanks for directing my paths to a relevant supervisor, Mr. Gordon Wayumba. It is through Mr. Wayumba's guidance and encouragement that this work came to completion. I also say thank you to Mr. Albert Inima, my second supervisor for his critiques that formulated the language that directed this work.

Thanks to Prof. Lawrence O. Gumbe, who believed in my abilities and whose prodding kept me going in a backdrop of financial and health challenges.

Special mention to Mr. Patrick Chebosi and the Mumias Sugar Company crew for helping with the field and laboratory work. To them, I owe much for the costs incurred to facilitate this work.

Mr. Christian Thine Omuto, who deserves more than a mention. He demystified script writing in ILWIS making the mysterious understandable. Mr. Onyango Mbunge did so much while proof reading the work.

Doreen and Emily for their love and family support.

To Mzee Odhiambo Oruoch, for laying in my youthful heart the vision to pursue academic excellence and my mom, Dorcas Atieno Odhiambo, for her prayers.

## TABLE OF CONTENTS

<b>DECLARATION</b> .....	<b>I</b>
<b>QUOTES</b> .....	<b>II</b>
<b>DEDICATION</b> .....	<b>III</b>
<b>ACKNOWLEDGMENTS</b> .....	<b>IV</b>
<b>LIST OF FIGURES</b> .....	<b>VIII</b>
<b>LIST OF TABLES</b> .....	<b>IX</b>
<b>ACRONYMS</b> .....	<b>IX</b>
<b>ABSTRACT</b> .....	<b>X</b>
<b>1. INTRODUCTION</b> .....	<b>1</b>
1.1 BACKGROUND INFORMATION .....	1
1.2 PROBLEM STATEMENT .....	4
1.3 JUSTIFICATION .....	5
1.4 RESEARCH QUESTION .....	6
1.5 OBJECTIVES OF THE STUDY .....	6
1.6 SCOPE OF STUDY .....	7
<b>2. LITERATURE REVIEW</b> .....	<b>9</b>
2.1 REMOTE SENSING TECHNOLOGY .....	9
2.1.1 Definition and description .....	9
2.1.2 Electromagnetic spectrum: transmittance, absorptance, and reflectance .	11
2.1.3 Satellite imagery and biomass quantification .....	12
2.1.4 Reflectance characteristics of green plants.....	13
2.1.5 Vegetation indices and their application .....	18
2.1.5.1 The normalized difference vegetation index (NDVI).....	19
2.1.5.2 Façade cap transformation index .....	22
2.1.5.3 The perpendicular vegetation index (PVI) .....	23
2.1.5.4 Soil Adjusted Vegetative Index (SAVI).....	24

2.1.6	Estimation of biophysical properties and agro-ecological indicators.....	24
2.1.6.1	Fractional vegetation cover .....	24
2.1.6.2	Leaf are index .....	25
2.1.6.3	Crop reflectance coefficient.....	25
2.1.6.4	Crop potential evapotranspiration .....	26
2.1.6.5	Fractional photosynthetically active radiation.....	26
2.1.6.6	Accumulated absorbed photosynthetic active radiation .....	27
2.1.6.7	Accumulated Biomass .....	28
2.2	SUITABILITY OF REMOTE SENSING FOR BIOMASS QUANTIFICATION.....	30
2.3	LIMITATIONS OF REMOTE SENSING .....	32
<b>3.</b>	<b>MATERIALS AND METHODS.....</b>	<b>33</b>
3.1	STUDY AREA .....	33
3.2	STUDY PROTOCOL .....	37
3.2.1	Selection, sampling design and measurement of field data.....	38
3.2.2	Selection and processing of spectral data .....	40
3.2.3	Statistical (Geostatistical) analysis .....	45
3.2.4	Spatial quantification and mapping of sugarcane biomass.....	49
<b>4.</b>	<b>RESULTS AND DISCUSSIONS.....</b>	<b>50</b>
4.1	FIELD AND LABORATORY DATA .....	50
4.2	CROP REFLECTANCE DATA EXTRACTION.....	50
4.3	QUANTIFYING SUGARCANE BIOMASS USING REGRESSION ANALYSIS .....	57
4.3.1	Biomass quantification model and mapping .....	69
4.3.2	Model validation.....	70
4.3.3	Errors and Contingencies.....	72
<b>5.</b>	<b>CONCLUSION AND RECOMMENDATIONS.....</b>	<b>75</b>
5.1	CONCLUSION .....	75
5.2	RECOMMENDATIONS .....	76
	<b>REFERENCES .....</b>	<b>77</b>
	<b>FURTHER BIBLIOGRAPHY .....</b>	<b>83</b>

<b>GLOSSARY OF TERMS .....</b>	<b>86</b>
<b>APPENDICES.....</b>	<b>87</b>
A.1    SCRIPT FOR ESTIMATING BIOPHYSICAL PARAMETERS .....	87
A.2    STATISTICAL ANALYSIS.....	91



## LIST OF FIGURES

Figure 1-1: Crop phenological phases (A-E).....	3
Figure 1-2: Four sampling plots within a sampling block.....	8
Figure 2-1: Energy interaction response,.....	12
Figure 2-2: General spectral response of green vegetation .....	15
Figure 2-3: Spectral signatures of water, soil and vegetation .....	17
Figure 3-1: Study area .....	34
Figure 3-2: Sugar production by factory in Kenya. ....	35
Figure 3-3: Sugarcane fields in Mumias Nucleus Estate .....	36
Figure 3-4: W-shaped sampling transect.....	39
Figure 3-5: The Global Positioning System handsets and a network of satellites .....	41
Figure 4-1: Geo-referenced image clipped to show plots 34 and 35.....	51
Figure 4-2: Spectral signatures of sugarcane canopy in Fields 34 and 35 .....	55
Figure 4-3: Spectral response of Mumias sugarcane canopy at senescence .....	56
Figure 4-4: Correlation coefficients for band values (bi-pixel mean values).....	58
Figure 4-5: Correlation coefficients for band values (pixel by pixel mean values) .....	59
Figure 4-6: ETM+ reflectance vs measured biomass (pixel by pixel) on Field 35 .....	60
Figure 4-7: ETM+ reflectance vs. measured biomass: all sampling plots (Field 35)...	63
Figure 4-8: Biophysical parameters vs biomass: all sampling plots (Field 35).....	64
Figure 4-9: Biophysical vs measured biomass pixel by pixel (Field 35) .....	65
Figure 4-10: Band reflectance vs. measured biomass: all sampling plots (Field 34)...	66
Figure 4-11: Biophysical vs measured biomass bi-pixel by bi-pixel (Field 34) .....	67
Figure 4-12: Biophysical parameters vs biomass: all sampling plots (Field 34).....	68
Figure 4-13: Predicted sugarcane biomass for Mumias Nucleus Estate .....	70
Figure 4-14: Comparison between measured and estimated biomass (Field 35).....	71
Figure 5-1: Sugarcane fields (harvested and growing).....	76

## LIST OF TABLES

Table 3-1: The spectral bands: LANDSATETM+ satellite .....	44
Table 4-1: Measured Biomass data for Fields 34 and 35 at the sampling plots.....	50
Table 4-2: Spectral analysis and biophysical estimation for Field 34.....	52
Table 4-3: Spectral analysis and biophysical estimation for Field 35.....	53
Table 4-4: Coefficient of correlation values ( $r^2$ ) .....	57
Table 4-5: Performance of prediction and validation models .....	73
Table 4-6: Jack-knife cross validation results .....	74

## ACRONYMS

ETM	Enhanced Thematic Mapper
GIS	Geographical Information Systems
GPS	Global Positioning Systems
NDVI	Normalized Vegetative Index
RS	Remote Sensing
SAVI	Soil Adjusted Vegetative Index
TM	Thematic Mapper
TSAVI	Transformed Soil Adjusted Vegetative Indices
VI	Vegetative Index

## ABSTRACT

The massive use of imported fossil fuels is a major cause of concern in Kenya. Energy from biomass could be used to address this problem. In the energy sector, sugarcane biomass can provide a safe, sustainable and potentially big source of energy that could trigger and sustain the industrialization process in Kenya. A sustainable biomass system forms a closed cycle of carbon increasing the importance of fuel from biomass in terms of global and regional environments. However, to tap this resource, requires the establishment of the temporal and spatial variability of the biomass distribution. There is also need to accurately determine the total amount of biomass that is available for use by the energy conversion systems in Kenya.

This study attempted to provide part of the answer to the question of spatial and temporal variability of biomass by developing a technique for quantifying and mapping sugarcane biomass using remote sensing in sugarcane fields within the Mumias Nucleus Estate. The technique developed aimed at providing a methodology for quantifying sugarcane biomass in real time to reduce inaccuracies, high costs and tediousness associated with the current physical crop cut census techniques.

The study was done and point at the existence of relationships between spectral reflectance measurements and leaf area index, vegetation cover and total aboveground green biomass. Pertinent crop biophysical parameters that included leaf area index, fractional photosynthetic active radiation and leaf cover that characterize biomass in sugarcane fields were estimated from satellite image while spectral reflectance values were used in band combinations to calculate vegetative indices. Five vegetation indices [Normalized Difference Vegetation Index, Soil Adjusted Vegetative Index, Red/Near Infrared, Red/Middle infrared and Near Infrared/Middle infrared] were calculated and compared with in-situ ground truthed data clipped in two fields in Mumias Nucleus Estate.

Results indicated that biomass production and yield of a sugarcane crop could be quantified by satellite imagery by employing any of the vegetation indices

investigated. The biophysical parameters (Leaf area index, fractional cover, and fractional photosynthetic active radiation) were also good quantifiers. Leaf area index and Red/Near infrared were the best biomass quantifiers with coefficients of correlation ( $r^2$ ) of 0.94 each. There also existed a strong relationship between the spectral values and field biomass with the best prediction having correlation coefficients ( $r^2$ ) of 0.78 and 0.82 for band 3 and band 4 respectively. Temporal maps of vegetation indices developed using transformed values of band 3 and band 4 suggested that yield and biomass could be determined by Enhanced Thematic Mapper (ETM+) satellite imagery over the senescence stage of sugarcane crop growth. The model developed performed well giving a coefficient of efficiency ( $C_{eff}$ ) between estimated and measured yields of 0.98 and an average relative percent error of 0.0175. The study showed that remote sensing provides fast data gathering over large areas and that these data sets may be used to estimate the final crop yield. This allows for practicing timely field management operations to yield, which is one of the major elements in biomass utilization.

## 1. Introduction

### 1.1 Background Information

Monitoring biophysical features of sugarcane using remotely sensed data is crucial for many purposes, such as estimating productivity of growing and mature cane, sugarcane health and vigor monitoring, and understanding sugarcane ecological processes. Studies have been done and point at the existence of relationships between spectral reflectance measurements and leaf area index (LAI) and total aboveground green biomass (Guo *et al.*, 2000).

In the energy sector, sugarcane biomass provides a safe, sustainable and potentially big source of energy that can trigger the industrialization process in Kenya. A sustainable biomass system forms a closed cycle of carbon increasing the importance of fuel from biomass in terms of global and regional scales. A huge potential exists for utilizing agricultural waste and energy crops to power irrigation, industrial and production plants in Kenya. A rough estimate shows that more than 3.6 million tones of biomass were available in 1986 for utilization in energy conversion systems in Kenya (Hankins, 1989). Out of this, bagasse accounted for 2.3 million tones of which 1.06 million tones were being utilized to power part of the production process (Hankins, 1989). Matsuda *et al.* (1982) pointed out that the potential contribution of biomass energy to the total energy consumption in the world is 16.6% with some regions consuming slightly larger amounts. The weight of biomass energy in developing countries is given as 38% (55% for Africa, Woods and Hall, 1994).

However, to tap this resource, requires the knowledge and establishment of the temporal and spatial variability of the biomass distribution. There is also need to accurately determine the total amount of biomass that is available for use in energy conversion systems in Kenya.

The study of the spatial and temporal distribution of the biomass to determine the biomass waste and that quantifies the available energy for production was necessitated

by this need. The study developed a methodology for quantifying and characterizing the biomass in temporal and spatial terms in a timely and reliable way and therefore helps to address the anomaly. The study used Remote Sensing (RS) techniques to provide a graphical and an analytical tool for mapping and quantifying biomass within the sugarcane zone. Remote sensing had the capacity to provide valuable information regarding various natural resources (Skidmore *et al.*, 1997 and Warren *et al.*, 1990).

Leaf area index (LAI) (established from the number and area of leaves) and, percent vegetation cover (%), species richness and unit area plant biomass or total aboveground green biomass ( $\text{g/m}^2$ ), canopy height (m) and plant moisture content ( $\text{g/m}^2$ ) are some of the biophysical parameters commonly used as indicators of the status of a sugarcane crop (Guo *et al.*, 2000, Maas, 1998, Wiegand *et al.*, 1992). The other incidental biochemical properties that characterize sugarcane are; the chlorophyll content and rate of photosynthesis, the brix, sucrose, purity and fiber content of the crop.

These biophysical and biochemical properties vary with the phenological development of the crop and therefore, their characteristics could be used individually or collectively to quantify crop biomass, yield and even status. The typical life cycle of an individual sugarcane tiller has been well described (Dillewijn, 1952) and can be used to identify different crop development phases ( Figure 1-1). The individual sugarcane tiller passes through different phenological stages distinguished by the level of tiller density (TD) expressed in tillers/ha. The TD, for well-watered crops, is normally influenced by three internal processes;

i. Primary tiller germination and emergence (A, B and C)

The first phase is the Pre-germination phase (A) that covers crop initiation up to the germination of the first bud. This part of the phase starts when the crop is either planted or when the previous ratoon is harvested. The buds remain dormant for a thermal period of 350 degree days (with base temperature of 9 ° C) for plant crops and 100°C days for ratoon crops.

The Pre-emergence phase follows with bud germination and continues until the emergence of the first primary tiller. After germination, the new shoot elongates towards the surface of the soil emerging as a primary tiller. The pre-emergence phase accounts for the time lag between germination and emergence. The primary tiller emergence phase requires a thermal duration of 170 degree days (at a base temperature of 9°C, where, base temperature ( $T_b$ ) is defined as the temperature below which the rate of development is zero.).

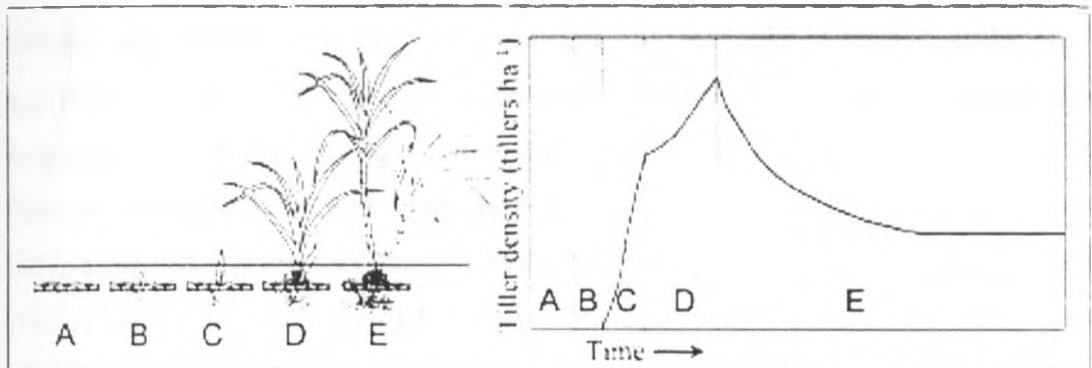


Figure 1-1: Crop phenological phases (A-E) and their relative influence on tiller density

ii. Underground branching, also known as tillering (D)

After emergence, leaves will develop and the primary tiller will start tillering. The tillering process produces a stool of upright stalks containing one primary tiller and various numbers of higher order secondary and tertiary tillers. Dillewijn (1952) identified light intensity and day length as the most important driving factors for tillering, while temperature was considered to be the second most important factor. There are two fairly distinct sub-phases during the tillering phase. These are;

- An early sub-phase of profuse tillering and
- A sub-phase with a decline in tillering and more distinctive stalk elongation.

This occurs due to light competition at a base temperature of 16°C (Inman-Bamber, 1994). This constant thermal period is termed the telomechron interval derived from the word telome which refers to a terminal branch of a vascular plant (Gray, 1967). The branching behaviour of tillers implies an exponential increase in tiller density over thermal time.

Inman-Bamber (1994) demonstrated that the stalk population of sugarcane at the tiller emergence phenological stage was highly correlated with thermal time when using a base temperature ( $T_b$ ) of  $16^{\circ}\text{C}$  and that the stage spans from 0 to  $500^{\circ}\text{C}$  days, at which the stalk population peaks. The development of the sugarcane ratoon canopy may be viewed as a process dependent on the emergence of tillers from the soil, and leaves from the whorl of each tiller and therefore falls in this phase.

Tiller emergence is of prime concern in the detection of sugarcane in remote sensing applications that rely on measurements of light reflected from the sugarcane canopy. It is therefore a suitable phase to use to derive biophysical parameters of sugarcane using remotely sensed data.

iii. Tiller senescence due to light competition (E).

Tiller senescence phase occurs from the period of the first tiller senescence to the time of harvest which is normally between  $500$  and  $1000^{\circ}\text{C}$  days, where tillers die off, followed by a relatively stable stalk population after about  $1600^{\circ}\text{C}$  days. Most of the sucrose accumulation occurs at these latter stages prior to harvesting

iv. Flowering

Sugarcane flowering depends on favorable temperature, sunlight and day length conditions, and does not always occur

The quality of a sugarcane crop, however, is described by its suitability for a defined purpose either for ethanol, fiber, and sugar or jaggery production

## 1.2 Problem Statement

Biomass energy supply is dependent on the spatial and temporal variability of the supply of the material. This therefore requires timely and accurate forecasts and predictions of the seasonal variability in supply for any sustainable and reliable harnessing to be applied for industrial applications.

Currently, the crop cut census method used to forecast and predict sugarcane biomass production is tedious, inaccurate and time consuming calling for the development of a



more appropriate methodology to ameliorate these problems. The development of a quantification tool based on remote sensing is therefore timely and relevant.

This study attempted to provide part of the answer to the question of spatial and temporal variability of biomass by using a rapid assessment technique to predict the availability of biomass in sugarcane fields within the Mumias Nucleus Estate

### 1.3 Justification

The massive use of imported fossil fuels is a major cause of concern in Kenya. Imported energy makes the country politically vulnerable, and drains the country's meager foreign exchange. Energy from biomass can be used to address this problem. In Finland, biomass accounts for 19% of the total energy requirements (Nyang, 1998). In Kenya, the most widely used fuels by the household consumers (which accounts for 58% of the national energy consumption) are Kerosene, firewood and charcoal with national penetrations of 94%, 74% and 57% respectively (Nyang, 1998). Only 5% of the rural population has access to electricity while 14% of households in the country use electricity. Among rural users of electricity, 52% obtain their supplies from the national grid, 33% from solar panels and 15% from diesel or petrol powered generators (Nyang, 1998).

Kituyi *et al.* (2001a), approximates the per capita annual consumption of fuel-wood in Kenya as 690-890 kg for firewood and 70-110 kg for charcoal while Nyang (1998) places it at 770 kg for firewood and 1090 kg for charcoal. Official reports for Kenya place the demand for energy from biomass at 68% of the total energy supply (Kituyi *et al.*, 2001b). The importance of this resource is therefore clear. The discrepancies in the estimates result from the methods employed in carrying out the baseline survey. Use of remote sensing, GIS and GPS techniques to make estimates could address this anomaly.

Effective biomass resource management for any production system is dependent on accurate assessment of the kind, amount, and condition of existing biomass resources.

Kituyi *et al.* (2001a) ascertains that the temporal and spatial patterns of bio fuel in Kenya are not known.

Kenya depends significantly on agriculture for its economic growth with 26% of the GDP coming from this sector compared to 14% from industry, and 60% from services. With a population of 28.6 million people (Kituyi *et al.*, 2001b), 77% of these people live within the rural area and depend directly on the natural resource base for their livelihoods. The country is a biomass based subsistence economy because of the direct dependence of her rural people on plant and animal products (Nyang, 1998)). The study area covered an important resource and segment of the country's economy.

#### **1.4 Research Question**

Can remote sensing techniques provide an effective system for quantifying sugarcane biomass in real time and thereby provide a management decision tool for sugarcane biomass utilization?

#### **1.5 Objectives of the study**

Overall objective

The broad objective of the study was to develop a methodology for quantifying biomass in sugarcane fields using Remote Sensing (RS) techniques.

Specific objectives

The specific objectives of the study were to;

- 1) Identify the pertinent biophysical parameters, which characterize biomass in sugarcane fields.
- 2) Use satellite data to develop a methodology for quantifying the parameters identified in (1) above.
- 3) To determine parameters identified in (1) above using the methodology developed in (2) above.

- 4) Using the parameters identified in (1) above and the quantities determined in (3) above to develop and solve a mathematical model for spatial quantification of sugarcane biomass in the fields.
- 5) Using ground-truthed field data to verify the solution in (4) above, and integrate the verified solution to map biomass in sugarcane fields in dry matter (tones/ha) using remotely sensed data.

## 1.6 Scope of study

The study focused on establishing an empirical relationship between the remotely derived biophysical parameters and vegetative indices (VI) of sugarcane and ground observations for spatial quantification of sugarcane biomass in Mumias Nucleus Estate. The terms biophysical parameters as used in this study imply leaf area index (LAI), above ground cover (fcover), and fractional photosynthetic active radiation (fPAR). The vegetative indices and band ratios used for quantification in this study were; Normalized Difference Vegetative Index (NDVI), Soil Adjusted vegetative Index (SAVI), Band3/Band4 index (R/NIR), Band3/Band6 index (R/MIR) and Band4/Band6 index (NIR/MIR).

Two fields with sugarcane at fifteen and sixteen months of growth were used for the study. Each field was randomly divided into four blocks each measuring 20 m by 20 m. Observations were made in a cross (⊕) shaped transect at four sampling plots of 1.5 m by 1.5 m sizes in each randomized block (

Figure 1-2). Although ground measurements were of a high spatial resolution, the results were extrapolated to half a pixel and a full pixel to conform to the satellite spatial resolution of 30 m by 30 m of the Enhanced Thematic Mapper (ETM+) used.

Biomass weight in kg was measured using a weighing scale while the clipped plots were geo-referenced using a GARMIN® global positioning handset (GARMIN, 1998). Linear regression statistics were used to establish the relationship between biophysical parameters and the observed above ground biomass. The relationship derived was then used to select the most pertinent biophysical parameter that most

efficiently quantified sugarcane biomass. This parameter was then used to quantify biomass in dry matter terms from remotely sensed ETM+ 2003 February image clipped to Mumias Nucleus Estate boundaries and to make a spatial representation through a script written and run in ILWIS 3.2 Academic ®.

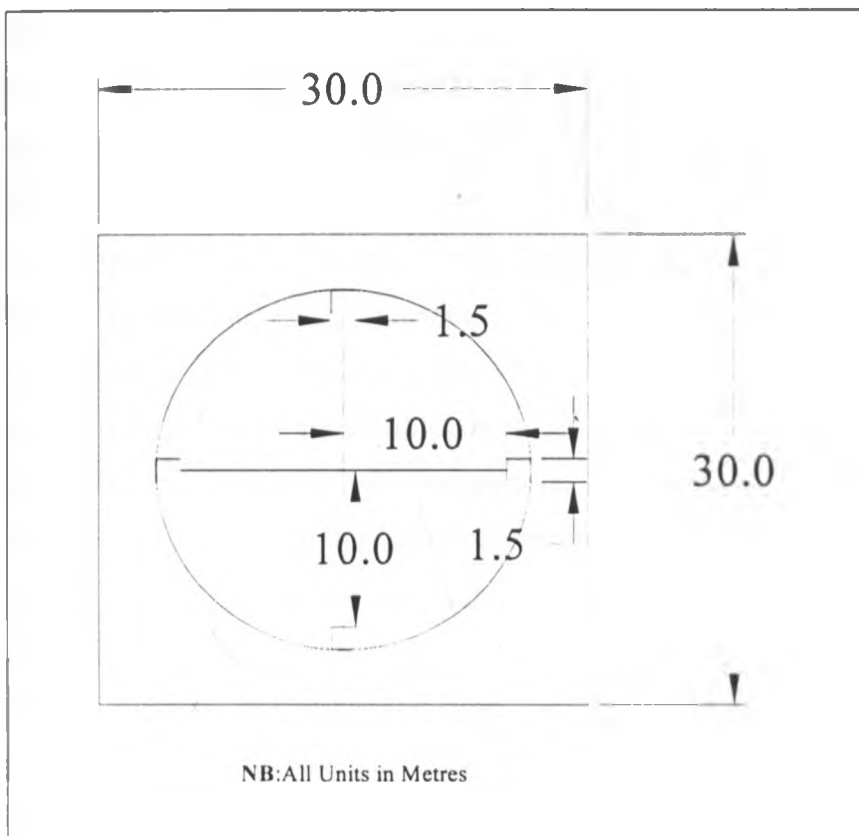


Figure 1-2: Four sampling plots within a sampling block.

## 2. Literature Review

### 2.1 Remote Sensing Technology

#### 2.1.1 Definition and description

Remote sensing is the science and art of obtaining information about an object, area or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area or phenomenon under investigation (Lillesand and Kiefer 2000; Woldu, 1997). A 1967 report defined remote sensing as “any means of gaining information without direct contact (Hardy, 1989). Other definitions qualify the means of observation and limit it to electromagnetic waves of the optical and microwave range. The restriction to electromagnetic waves is due to the fact that the observation from spacecraft excludes other possibilities such as sonic wave, which require a medium like air, water or solid earth for propagation (Hartl, 1989). Other means of indirect observation by, for example, stationary magnetic or electric fields are not sensitive enough for high geometric resolution measurements. Remote sensing is characterized by its synoptic coverage of the earth’s surface at different spatial and spectral resolution and with different types of sensors. It has the capacity for repeated coverage’s and temporal analysis and measures qualitatively and quantitatively features of the earth’s surface without material contact. Two types of sensors could be used. “Active sensors” if they consist of both a transmitter and a receiver or “passive sensors” where only receivers are used as is the case with external signal sources (Hartl, 1989).

The targets’ features can be measured indirectly through the interaction with the electromagnetic waves. Various methods are applicable. In the passive case one determines any of the following:

- The amount of electromagnetic emission of the target itself; the reflectivity of the optical signal power of the sun at the target; the amount of scattered power, if the “target” is, for example, an aerosol of the atmosphere; the absorption of solar radiation power in the media to be measured.

- Other methods can also be applied, for example, the determination of the amount of signal polarization or depolarization. In all cases it is important to identify functional dependencies of the features measured on frequency or angle of observation.

In the active system, it is the instrument itself that is the source of radiation. It sends signals to target under investigation, and receives the return signal, having the unique characteristics of the target features (Woldu, 1997).

As far as satellite remote sensing is concerned, systems operating in the visible and infrared part of the electromagnetic spectrum are passive while the microwave instruments are either active or passive.

Currently, satellites are the main devices in remote sensing. The two main types of satellites are the sun-synchronous and geostationary satellites. The sun-synchronous satellites keep a precise pace with the sun's westward progress as the earth rotates so that they always cross the equator at precisely the same solar time. Examples are LANDSAT and NOAA satellites.

Geostationary satellites spin on their own orbit and are placed around the celestial equatorial plane at an altitude of 36,000 km above the earth's surface maintaining a period of 24 hours by spinning with an angular velocity that matches the earth's rotation. As a result, they remain at the same point above the earth at all times appearing to be stationary. An example is Meteosat. These satellites are helpful in obtaining constant and persistent image of a particular area at fixed interval that is a great advantage in monitoring a location with high temporal resolution to capture the transient behavior of objects such as rain clouds.

### 2.1.2 Electromagnetic spectrum: transmittance, absorptance, and reflectance

The primary source of energy that illuminates natural targets is the Sun. Solar irradiation (also called insolation) arrives at earth at wavelengths which are determined by the photospheric temperature of the sun (peaking near 5600 °C). The main wavelength interval is between 200 and 3400 nm (0.2 and 3.4  $\mu\text{m}$ ), with the maximum power input close to 480 nm (0.48  $\mu\text{m}$ ), which is in the visible green region. As solar rays arrive at the earth, the atmosphere absorbs or backscatters a fraction of them and transmits the remainder.

Upon striking the land and ocean surface (and objects thereon), and atmospheric targets, such as air, moisture, and clouds, the incoming radiation (irradiance) partitions into three modes of energy-interaction response (Figure 2-1):

- i. Transmittance ( $\tau$ ) - some fraction (up to 100%) of the radiation penetrates into certain surface materials such as water and if the material is transparent and thin in one dimension, normally passes through, generally with some diminution.
- ii. Absorptance ( $\alpha$ ) - some radiation is absorbed through electron or molecular reactions within the medium ; a portion of this energy is then re-emitted, usually at longer wavelengths, and some of it remains and heats the target;
- iii. Reflectance ( $\rho$ ) - some radiation (commonly 100%) reflects (moves away from the target) at specific angles and/or scatters away from the target at various angles, depending on the surface roughness and the angle of incidence of the rays.

Because they involve ratios (to irradiance), these three parameters are dimensionless numbers (between 0 and 1), but are commonly expressed as percentages, Following the Law of Conservation of Energy:  $\tau + \alpha + \rho = 1$  (Lillesand and Kiefer, 2000).

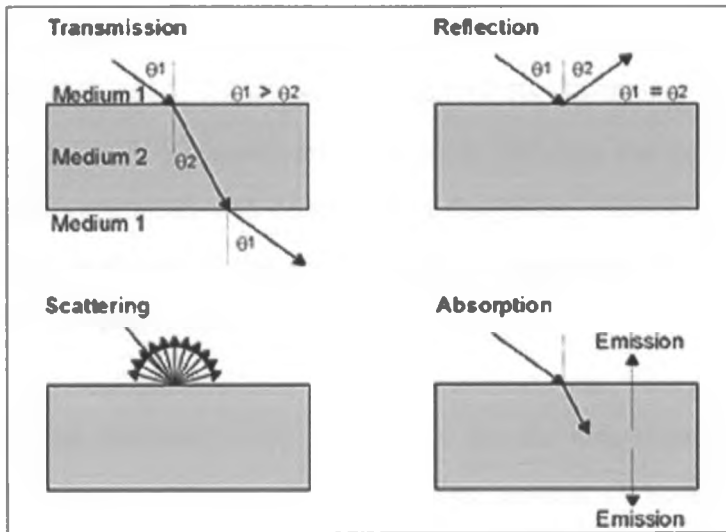


Figure 2-1: Energy interaction response, *Source: Lillesand and Kiefer, 2000*

A fourth situation, when the emitted radiation results from internal atomic/molecular excitation, usually related to the heat state of a body, is a thermal process.

Remote sensing of the earth traditionally has used reflected energy in the visible and infrared and emitted energy in the thermal infrared and microwave regions to gather radiation that can be analyzed numerically or used to generate images whose tonal variations represent different intensities of photons associated with a range of wavelengths that are received at the sensor.

### 2.1.3 Satellite imagery and biomass quantification

Since the successful launch of the Earth Resource Technology Satellite-1 (ERTS-1, renamed LANDSAT-1) in 1972, digital remote sensing has been used with some success to monitor natural resources and to provide input to better manage the earth Skidmore *et al.* (1997). The Advanced Very High Resolution Radiometer (AVHRR) imagery when used in time series can be applied to estimate global biomass production and also to relate the seasonal increase and senescence to agricultural production Skidmore *et al.* (1997). The NDVI derived from AVHRR data has been extensively used for vegetation monitoring, crop yield assessment and drought detection Albert *et al.* (2002). Researchers have found relationships between vegetation properties and remotely sensed variables. Of more significance is the fact



that vegetation productivity is more strongly related to TM band ratios than individual bands. Ratios of middle-infrared to near-infrared, as well as ratios of middle-infrared to near infrared and visible (blue), were found to be significant (Skidmore *et al.*, 1997). Skidmore *et al.* (1997) analyzed (simulated) TM data and concluded that most information about vegetation was contained in the green, near infrared and middle-infrared. Detailed analysis of the relationships is presented in Skidmore (1997), Lillesand and Kiefer (2000) and Jensen (1996)

Remote sensing has become a very useful tool for the classification of vegetation cover especially where fieldwork requires a large amount of economic and human resources (Betrann and Belmonte, 2001). In agriculture, there is need for regular information on the growth cycle of crops and their sequence in space and time. Agriculture is a dynamic system whose control and management call for rapid and regular acquisition of a great many data on the agronomists' familiar triad (Frayse, 1980). The identification of a crop and its spatial distribution over an area could be used to improve the management of the crop. The principal parameters could be used to estimate energy in biomass and the surface occupied by each class of crop each of which has a different energy value.

#### **2.1.4 Reflectance characteristics of green plants**

Reflectance of green plants is an interaction and integration of four things:

- The reflectance of plant parts (single parts, organs, pigments);
- the reflectance of sets (canopies);
- the nature and state of the plants and
- The structure and texture of the set (Frayse, 1980).

The biophysical properties of a crop can influence or cause spectral reflectance variation due to internal or external factors. Internal factors like age of organ, location of the organ (below or above the surface) and state of turgescence alter the plant reflectance properties through the changes they have on plant water content, turgescence, mesophyll structure, evapotranspiration and pigmentation. External

factors could be at the ground level (availability water, mineral ions and toxic salts), atmospheric level e.g. through climatic factors, biological pathogens and irradiance incident angle.

When incident solar radiation interacts with green leaves, a high level of absorption, approximately 80% to 90%, of the visible light by chlorophyll and carotenoid pigments takes place in the 0.3-0.4 $\mu\text{m}$  ( blue part) and 0.6  $\mu\text{m}$  ( red part) of the electromagnetic spectrum. But in the 1.3-2.5 $\mu\text{m}$  region, there is absorption due to the presence of liquid water. This phenomenon is due to the presence of active chloroplasts.

At 0.5  $\mu\text{m}$  and the 0.74–1.2 $\mu\text{m}$  region, there is high reflectance and transmittance of the near-infrared light from the leaves. At 0.5  $\mu\text{m}$  (green region), healthy green vegetation generally reflects 20% of the incident near-infrared energy while at 0.74–1.2 $\mu\text{m}$  region 60% of near-infrared light is reflected. In the 1.3-2.5 $\mu\text{m}$  region the presence of liquid water causes a general decrease of reflectance for mid-infrared bands. This is explained by the scattering of radiation caused by the leaf structure.

Most of the chloroplasts are present in the upper layers of leaves (the palisade cells just below the transparent cuticle and upper epidermis). Thus visible light is either absorbed (red and blue light) or reflected (green light, normally) from these layers in leafy canopies. Very little light penetrates below this layer (D'Souza *et al.*, (1995); Jensen, (1996)).

At the plant canopies, spectral absorption and reflection determine the spectral radiance of the soil-plant complex and this parameter provides the basis for passive remote sensing. The objective is thus to derive information on the nature and state of the vegetation by using specific wavelengths selected to provide a strong signal from the vegetation and a good spectral contrast with the background. Figure 2-2 shows several regions of the spectrum where such parameters can be evaluated for leafy and canopy green vegetation, respectively.

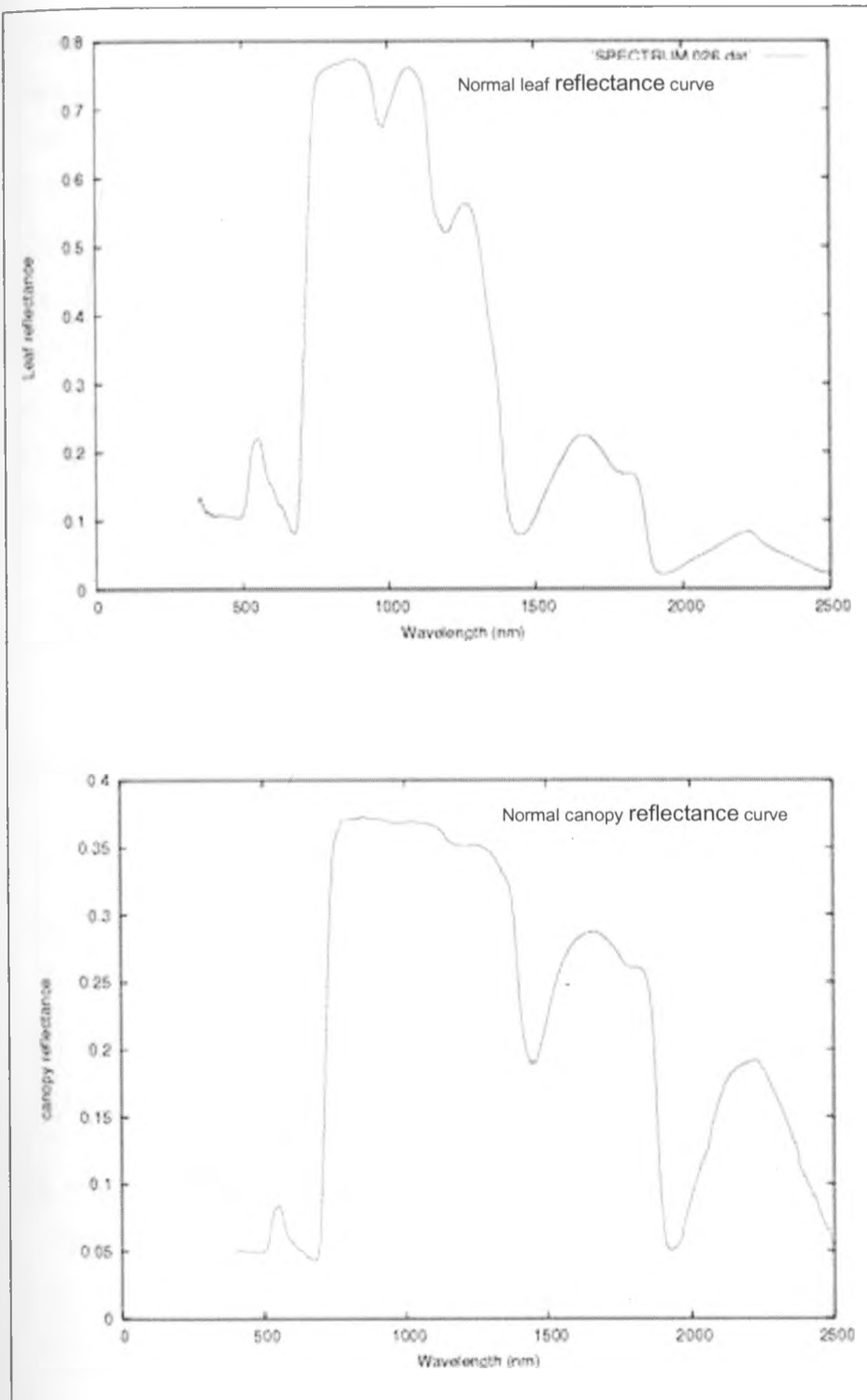


Figure 2-2: General spectral response of green vegetation Source: (D'Souza et al., 1995)

Parodi, (2000) gives a detailed description of how to determine the structural biophysical parameters of a crop using satellite imagery data. A strong correlation appears to exist between remotely acquired data and the concentration of various biochemicals within the vegetation canopy (Curran *et al.*, (1997). A significant correlation has been shown between crop chlorophyll concentration and spectral measurements (Patel *et al.*, 2001; Jago *et al.*, 1999 and Munden *et al.*, 1994). Inoue *et al.* (2000) used a dirigible-mounted camera to collect spectral measurements in four bands, allowing them to successfully estimate LAI and fresh biomass for soybean (*Glycine max*) and rice (*Oryza sativa*). Inoue and Morinaga (1995) had previously found strong correlations between remotely sensed parameters and fresh biomass ( $r = 0.97$ ) and greenness ( $r = 0.68$ ) using spectral observations. Brown *et al.* (1997) demonstrated that canola (*Brassica napus*) biomass and leaf area were significantly correlated to near-infrared (NIR) reflectance. Similarly, Cloutis (1999), using high spectral resolution optical and radar imagery, found several statistically significant correlations between spectral observations and LAI, plant height, and canopy temperature for different crops. Thenkabail *et al.* (2000) used a hand-held spectroradiometer to correlate biophysical parameters of cotton (*Gossypium hirsutum*), potato (*Solanum tuberosum*), soybean, maize, and sunflower (*Helianthus annuus*) with spectral observations. Several attempts have been made to use remote sensing technology to forecast crop yields for wheat (Asrar *et al.*, 1985), barley (*Hordeum vulgar*, Klemm *et al.*, 1987), soybean, and maize (Crist *et al.*, 1984 and Holben *et al.*, 1980).

The correlation between biophysical factors and spectral reflectance is influenced by width and location of bands within the electromagnetic spectrum. At the senescence, the chlorophyll breaks down which leads to a marked increase in reflectance, particularly of the red colors. Other pigments such as the orange carotenes and yellow xanthophylls, plus non-photosynthesis pigments such as red anthocyanins are often still present. No longer masked by the presence of chlorophylls, these pigments are largely seen in brown and yellow colors of senescent vegetation (D'Souza *et al.*,

1995). Conversely, dead or senescent vegetation reflects less than green vegetation in the reflective infrared region. Dry soil generally has higher reflectance than green vegetation and lower reflectance than dead vegetation in the visible region, whereas, in the near-infrared, dry soil generally has lower reflectance than green or senescent vegetation (Figure 2-3). Most vegetation indices are based on the fact that there are significant differences in the shape of these three curves. Remote sensing can provide useful information of the vegetation condition information if the three curves are placed on top of one another, Jensen (1996) (Figure 2-3).

As a plant canopy progresses through growth and development, from bare soil to a sparse canopy, to a fully developed canopy and then to partial or complete senescence into leafy litter, the proportion of red and blue light being absorbed by the varying amounts of leafy pigments, and the amounts of near-infrared light being reflected from developing leaf structure changes. Multi-spectral measurements taken through this sequence of change can therefore be used to monitor the phenological changes with time (D'Souza *et al.*, 1995).

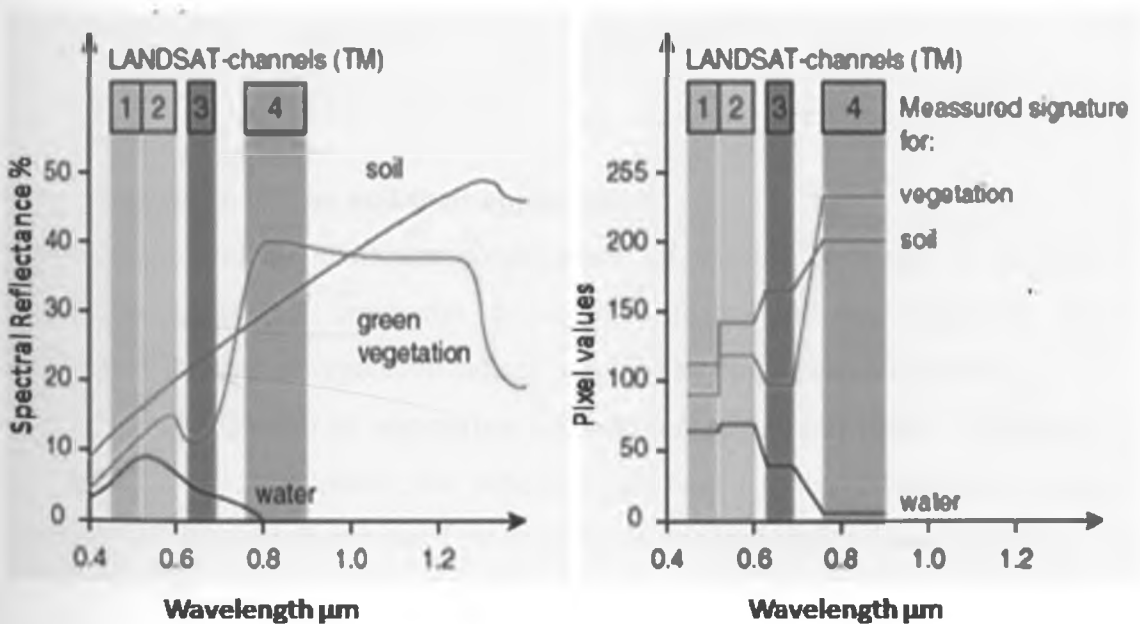


Figure 2-3: Spectral signatures of water, soil and vegetation source: ([www.geografforlaget.dk/course/ENGLISH/basic/tele1.htm](http://www.geografforlaget.dk/course/ENGLISH/basic/tele1.htm))

Temperature drives many of the processes responsible for phenological development in the crop. The effect of temperature is often quantified using the thermal time concept. Thermal time that is expressed in growing degree days (GDD) is defined as the cumulative value of the mean daily temperature minus the base temperature. The growing degree-day (GDD) values can be calculated by using equations 2-1 and 2-2 (Villa Nova *et al.*, 1972).

i) When  $T_m > T_b$  then;

$$GDD = \left( \frac{TM + T_m}{2} \right) - T_b \quad 2-1$$

ii) When  $T_m \leq T_b$

$$GDD = \left( \frac{(TM - T_m)^2}{2(TM - T_m)} \right) \quad 2-2$$

Where:

- GDD = Growing degree-days
- TM = Maximum daily air temperature
- Tm = Minimum daily air temperature
- Tb = Basal temperature

### 2.1.5 Vegetation indices and their application

Vegetation indices are particular combinations of spectral responses in different wavelength bands that emphasize a particular feature of the vegetation. The development and use of vegetation indices is guided by three general objectives:

- i. To enhance, through an appropriate combination of spectral bands, the relevant vegetation features. Ideally, the indices should have a better-defined relationship with physiological properties of the crop than individual spectral measurements.
- ii. To standardize the representation of the crop spectral response (Useful in region to region or year to year comparisons)
- iii. To reduce the dimensions of the data sets (index reduces to one data set than the sets which contribute to its calculation).

Much of the research in this area has involved the analysis of LANDSAT scanner (MSS) and Thematic data (TM) using digital image processing techniques. The goal has often been to reduce the much band information to a single number per pixel that predicts or assesses such canopy characteristics as biomass, productivity, leaf area index, and/or percent vegetation ground cover.

The AVHRR data on the other hand have been used extensively for large-area vegetation monitoring. Typically, the spectral bands used for this purpose have been the channel 1 visible band (0.58  $\mu\text{m}$  to 0.68  $\mu\text{m}$ ) and channel two near-IR band (0.73  $\mu\text{m}$  to 1.1  $\mu\text{m}$ ). Various mathematical combinations of the AVHRR channel 1 and 2 data have been found to be sensitive indicators of the presence and condition of green vegetation. These mathematical quantities and the algorithms used to extract information from remote sensed data have been referred to as vegetation indices (Jensen, 1996, Lillesand and Kiefer, 2000).

Vegetation index analysis must be based on a good knowledge of the spectral properties of vegetation as well as on the relationships between such spectral properties and plant growth (Malingreau, 1989). Malingreau (1989) gives a list of some of the vegetation indices used in remote sensing analysis as described in sections 2.1.5.1 to 2.1.5.4

*2.1.5.1 The normalized difference vegetation index (NDVI)*

There are several vegetation indices commonly referred to as Transformation Vegetation Index (TVI) or Normalized Difference Vegetation Index (NDVI). Most of these indices are ratios of various satellite bands, which are developed to accentuate the presence or absence of vegetation, thereby give an exposure of soil and rocks. The NDVI is defined as the ratio of the difference between the near infra red and the red reflectance to their sum (equation 2-3).

$$NDVI = \left( \frac{NIR - R}{NIR + R} \right) \quad 2-3$$

The NDVI values from different sensors can be calculated as shown in equations 2-4 to 2-6.

For LANDSAT MSS sensor

$$NDVI = \left( \frac{BAND7 - BAND5}{BAND7 + BAND5} \right) \quad 2-4$$

For LANDSAT ETM+ sensor:

$$NDVI = \left( \frac{TM4 - TM3}{TM4 + TM3} \right) \quad 2-5$$

For AVHRR sensor The NDVI is the difference of near-infrared (NIR, channel 2) and visible (VIS, channel 1) reflectance values normalized over the sum of channels 1 and 2 (equation 2-6). The NDVI equation produces values in the range of -1.0 to 1.0, where increasing positive values indicate increasing green vegetation and negative values indicate non-vegetated surface features such as water, barren, ice, and snow or clouds.

$$NDVI = \left( \frac{Ch2 - Ch1}{Ch2 + Ch1} \right) * k \quad 2-6$$

Where  $k$  is the factor for scaling the computed NDVI results from 16-bit to 8-bit range to minimize the volume and optimize analysis and display.

The vegetative index is called normalized because it is divided by the sum of radiances and thus normalizes somewhat for the difference in solar spectral irradiances. The NDVI is a measure of the amount and vigor of vegetation at the surface.

The magnitude of NDVI is related to the level of photosynthesis activity in the observed vegetation. In general, higher values of NDVI indicate greater vigor and amounts of vegetation (Parodi, 2000). NDVI is related to vegetation since healthy vegetation reflects very well in the near infrared part of the spectrum. Green leaves have a reflectance of 20% or less in the 0.5  $\mu\text{m}$  to 0.7  $\mu\text{m}$  range (green to red) and



about 60% in the 0.7  $\mu\text{m}$  to 1.3  $\mu\text{m}$  ranges (near infrared). The value is then normalized to  $-1 \leq \text{NDVI} \leq 1$  to partially account for differences in illumination and surface area.

Various studies have attempted to relate the performance of a crop to the vegetation index. Field measurements have established that there is a relationship between green biomass and the NDVI up to a certain value of the cover biomass above which the NDVI remains constant (saturation level). Brown *et al.* (1987) correlated the botanical indices with the mean reflectance values for all the seven LANDSAT TM bands and a NDVI value for the field for a 1988 LANDSAT TM image. The results indicated a significant positive relationship between the indices with bands 3 and 7, and a negative relationship with band 4 and NDVI.

Brightness values from individual MSS band (e.g., LANDSAT MSS4, MSS5, MSS6 and MSS7) have been used as vegetation indices to estimate percent ground cover and vegetative biomass. Correlation coefficients from 0.30 for MSS7 with crop cover to 0.88 for MSS6 with leaf area index have been reported. Between-band ratios of MSS and TM brightness values has been used to estimate and monitor biomass (Jensen, 1996).

Malingreau (1989) distinguished two approaches to the study of vegetative index:

- i. Study of NDVI at a specific time of crop development and
- ii. The study of the temporal evolution of NDVI of the crop performance. .

Malingreau (1989) recommends the use of an integration of the two NDVI study approaches for a more indicative result. Prince *et al.* (1991) derived an equation relating NDVI to biomass (Tree leaves and trunks) for a wide range of studies for three months of rainy season as:

$$\text{Biomass kg/ha/3 months rain} = 96 + 110(\text{NDVI}) \text{ days} \quad 2-7$$

He gave the relationship between NDVI and pixel brightness as;

$$\text{NDVI} = 0.003828P - 0.317311 \quad 2-8$$

### 2.1.5.2 *Façade cap transformation index*

This is based on the Gram-Schmidt sequential orthogonalization technique that produces an orthogonal transformation of the original four-channel MSS data-space to a new four-dimensional space (Jensen, 1996). The method has been rigorously tested and has been found quite useful in agricultural research. The transformation identifies four new axes that can be derived by equations 2-9 to 2-14;

- i. Soil Brightness index (SBI),
- ii. Green Vegetation Index (GVI),
- iii. Yellow Stuff Index (YVI) and
- iv. Non-Such Index (NSI), which is associated with atmospheric effects.

Generally, the first two indices (SBI and GVI) contain most of the scene information (95 to 98%). Jensen (1996) documents that, nearly all (98%) of the variance in bare soil spectra from several different soil types could be explained by the soil brightness index (SBI). Jensen (1996) concludes that bare soils would lie in a line that could be applied to LANDSAT MSS agricultural scenes. Greenness is an orthogonal deviation from the mean solid line and is used as a measure of the green vegetation present. The further the distance perpendicular to the soil line, the greater the amount of vegetation present within the field of view of the pixel. From their development of the vegetation indices, Jensen (1996) documents the use of LANDSAT MSS bands to derive the following equations:

$$\text{SBI} = 0.332 \text{ MSS4} + 0.603 \text{ MSS5} + 0.675 \text{ MSS6} + 0.262 \text{ MSS7} \quad 2-9$$

Where:

SBI= Soil Brightness index

Multiplying each of the four original LANDSAT MSS brightness values for each pixel by the corresponding façade (tasseled) cap coefficient creates greenness vegetation Index (GVI) image.

$$\text{GVI} = -0.283 \text{ MSS4} - 0.660 \text{ MSS5} + 0.577 \text{ MSS6} + 0.388 \text{ MSS7} \quad 2-10$$

Where;

GVI = Green Vegetation Index

$$GVI = -0.2728(TM1) - 0.2174(TM2) - 0.5508(TM3) + 0.7221(TM4) + 0.733(TM5) - 0.1648(TM7) \quad 2-11$$

$$YVI = -0.899 MSS4 + 0.428 MSS5 + 0.076 MSS6 - 0.041 MSS7 \quad 2-12$$

Where:

YVI= Yellow Stuff Index

$$NSI = -0.016 MSS4 + 0.131 MSS5 - 0.452 MSS6 + 0.882 MSS7 \quad 2-13$$

Where:

NSI= Non-Such Index

For the LANDSAT TM, the brightness index (BI) is given by equation 2-14 below.

$$BI = 0.2909(TM1) + 0.2493(TM2) + 0.4806(TM3) + 0.5568(TM4) + 0.4438(TM5) + 0.1706(TM7) \quad 2-14$$

Where:

BI= Brightness Index

### 2.1.5.3 The perpendicular vegetation index (PVI)

Since the NDVI is not independent of the substrate, it presents disadvantages in situations of very low vegetation cover. The perpendicular distance to the "soil line" indicates plant development and was developed by Richardson and Weigand (Jensen, 1996). The soil line is first established in the near infrared/red plane for a given soil type. The presence of vegetation causes the points to depart from the local soil line. The magnitude of the displacement with respect to the soil line is related to the amount of vegetation present (Malingreau 1989).

$$PVI = \left[ (IR_{soil} - IR_{veg})^2 + (R_{soil} - R_{veg})^2 \right]^{1/2} \quad 2-15$$

Where  $IR_{soil}$ ,  $IR_{veg}$ ,  $R_{soil}$ ,  $R_{veg}$  are the coordinates of the vegetation and soil points respectively.

#### 2.1.5.4 Soil Adjusted Vegetative Index (SAVI)

SAVI is the Soil Adjusted Vegetative Index, which was introduced by Huete (1988). This index attempts to be a hybrid between the ratio-based indices and the perpendicular indices.

$$SAVI = \left( \frac{(1 + L)(Ch1 - Ch2)}{Ch2 + Ch1 + L} \right) \quad 2-16$$

Where;

“L” is non-dimensional correction factor which ranges from 0 for very high vegetation cover to 1 for very low vegetative cover. The most typically used value is 0.5 for intermediate vegetative cover.

### 2.1.6 Estimation of biophysical properties and agro-ecological indicators

#### 2.1.6.1 Fractional vegetation cover

Because non-vegetated areas have no chlorophyll, separating non-vegetated, partially vegetated and densely vegetated land surface becomes possible after combining visible and near-infrared multispectral measurements (Parodi, 2000). Huete *et al.* (1985) showed a high correlation between vegetation cover and both spectral reflectance ( $r^2 = 0.90$ ) and NDVI ( $r^2 = 0.79$ ). Batchly *et al.* (1994) found that the SAVI concept was a better estimator of fractional coverage changes than NDVI-based time series. One way to normalize vegetation cover ( $V_c$ ) computations is to relate the values for soil without cover (SAVI) and for sparse canopies ( $SAVI_s$ ) to the values of dense canopies ( $SAVI_d$ ) (Equation 2-17) as  $SAVI_d$  and  $SAVI_s$  change with crop and soil type (Parodi, 2000). The equation applies to a single pixel and is more representative when a single crop species is encompassed by a single pixel. Equation 2-17 can also be applied with surface temperature and surface albedo to give a more favorable result as the near-infrared reflectance can both increase and decrease with fractional vegetation cover.

$$V_c = \frac{(SAVI - SAVI_s)}{(SAVI_d - SAVI_s)} \quad 2-17$$

Where;

- SAVI is the SAVI value for the current pixel (SAVI)

- $SAVI_s$  is the value of SAVI for soils without vegetation selected from the SAVI image
- $SAVI_d$  is the value of SAVI for dense canopies selected from the SAVI.

### 2.1.6.2 Leaf area index

Leaf area index (LAI) is the ratio of the total area of all the leaves on a plant to the area of ground covered by the plant. LAI represents the total biomass and is indicative of the crop yield, canopy resistance and heat flux. For most plants LAI increases with age (up to the beginning of senescence) and reaches a maximum value of 2.0 to 5.0.

Derived relationship between LAI and SAVI has been found to be linear during crop development stage until a threshold value of LAI (1.5 to 2) is reached. Choudhury (1994) simulated relationships between SAVI and LAI for different crops as:

$$LAI = -\frac{1}{c_3} \ln\left(\frac{c_1 - SAVI}{c_2}\right) \quad 2-18$$

The default values for  $c_1$ ,  $c_2$  and  $c_3$  are the average of many experiences developed by many researchers given as:  $c_1 = 0.69$ ,  $c_2 = 0.59$  and  $c_3 = 0.91$ .

### 2.1.6.3 Crop reflectance coefficient

The ground coverage detected from remote sensing for vegetation indexes gives the indication of stage of development of the crop and therefore the number of degree days after planting. This concept is used to derive the crop reflectance coefficient given as:

$$K_c = c_1 SAVI + c_2 \quad 2-19$$

The coefficients  $c_1$ ,  $c_2$  are crop dependent. Previous works by Neale *et al.* (1989) and Bausch (1993) have indicated that the reflectance based crop coefficient for corn responded to corn growth anomalies.

#### 2.1.6.4 Crop potential evapotranspiration

The crop potential evapotranspiration (ET<sub>c</sub>) is often calculated from meteorological data collected under conditions in which the actual evapotranspiration rate is less than the potential rate. Simplification of the Penman-Montieth method as proposed by Priestly and Taylor (1972) estimates evapotranspiration from well-watered surfaces through:

$$ET_{C-24} = \alpha \frac{\Delta}{\Delta + \gamma} (R_n - G) \quad 2-20$$

Where:

- ET<sub>c</sub> = Crop potential evapotranspiration
- $\alpha$  = the Priestly –Taylor coefficient.
- $\Delta$  = Slope of saturation vapor pressure [k Pa°C<sup>-1</sup>]
- $\gamma$  = Psychrometric constant [k Pa°C<sup>-1</sup>].
- G = Soil heat flux [W m<sup>-2</sup>]
- R<sub>n</sub> = Net radiation flux [W m<sup>-2</sup>].

The Priestly–Taylor method requires energy fluxes estimation of net radiation and soil heat flux.

#### 2.1.6.5 Fractional photosynthetically active radiation

Photosynthetically Active Radiation (PAR) describes the solar radiation available for photosynthesis. The fraction of solar radiation absorbed by chlorophyll pigments (fPAR) describes the energy related to carbon dioxide assimilation and is derived from PAR absorbed by canopy divided by PAR available from solar radiation. fPAR regulates the rate of carbon dioxide flow into the leaves. Asrar *et al.* (1992) stated that there is a definite experimental evidence that fPAR is monotonic and a near-linear function of the NDVI. The relationship of fPAR to NDVI, however, has been found to remain sensitive to the reflection properties of the background soil and litter (Hall *et al.*, 1990). This reason has been given to support the argument that postulates that Transformed Soil Adjusted Vegetation Index (TSAVI) and SAVI are better related to fPAR than NDVI. Asrar *et al.* (1992) modeled the relationship between fPAR and NDVI as:

$$fpar = offset + gain * NDVI$$

2-21

Where:

- NDVI is the normalized vegetation index map
- Values of gain = 1.2222 and offset = -1914

#### 2.1.6.6 Accumulated absorbed photosynthetic active radiation

This is the fraction of Photosynthetic Active Radiation (PAR) absorbed by the canopy that is used for carbon dioxide assimilation. AcAPAR is a measure for the accumulation of absorbed PAR (APAR) with time. PAR is normally part of the short wave solar radiation (0.3 to 3.0  $\mu\text{m}$ ) that supports photosynthesis in green plants. It relates to the visible part of the spectrum between 0.4 and 0.7  $\mu\text{m}$  where chlorophyll absorbs solar radiation. PAR is a fraction of the incoming solar radiation  $K \downarrow_{day}$  and varies with visibility, optical depth, and ozone amounts etc. (Frouin and Pinker, 1995). Values of AcAPAR can be estimated by using equation 2-25 derived as follows.

Theoretical estimation of PAR is given as:

$$PAR = c_1 * K \downarrow_{day} \quad 2-22$$

Where;  $c_1$  can be assigned a default value of 0.48.

However, only a fraction of the PAR is absorbed by the canopy and used for carbon dioxide assimilation resulting in APAR giving a leaf radiation balance of:

$$APAR = PAR - \rho_{PAR} * PAR - PAR_{trans} + PAR_{soil} \quad 2-23$$

Where;

- $\rho_{PAR}$  is the canopy reflectance at the upper side of the canopy in the 0.4 to 0.7  $\mu\text{m}$  spectral range,
- APAR is absorbed photosynthetic active radiation
- $PAR_{trans}$  Is the total amount of PAR that is transmitted through the canopy and directed to the soil.,

- $PAR_{soil}$  Is the portion of  $PAR_{trans}$  that is reflected from the soil underneath the canopy and is received back at the lower side of the canopy.

APAR can be deduced directly from PAR after simplifying equation 2-23 into:

$$APAR = fPAR * PAR \quad 2-24$$

Then:

$$AcAPAR = 10^{-6} * \sum APAR_i * t_i = \sum_{i=1}^n AcAPAR_i \quad 2-25$$

Where:

- PAR is the Photosynthetic Active Radiation,
- fPAR is the Fractional Photosynthetic Active Radiation.
- "t<sub>i</sub>" is the period over which the accumulation takes place for the current image (there are 'n' images in total) [seconds]. "APAR<sub>i</sub>" the APAR image that corresponds to a certain "t<sub>i</sub>" period.

#### 2.1.6.7 Accumulated Biomass

Biomass is living or recently living plant or animal matter. It can also refer to any particular part of a plant or organism as well. Defined as energy, the biomass resource can be considered as organic matter in which the energy of sunlight is stored in chemical bonds. When the bonds between adjacent carbon, hydrogen, and oxygen molecules are broken by digestion, combustion or decomposition these substances release stored energy.

Accumulated biomass is calculated as follows.

$$B_{act} = \sum_{i=1}^n \varepsilon(t) * AcAPAR_i \quad 2-26$$

Where:

- $\varepsilon$  (Several denominations) is the growth efficiency, conversion efficiency, biological efficiency, light use efficiency or radiation use efficiency [g



M/j]. It refers to the efficiency by which the APAR is converted to plant biomass increment.

- $AcAPAR_i$  is the accumulated absorbed photosynthetical active radiation (AcAPAR) for a certain time period.

Field *et al.* (1995) developed a global ecology model for net primary production in which  $\epsilon$  is calculated as:

$$\epsilon = \epsilon * T_1 * T_2 * W \quad 2-27$$

Where:

- $\epsilon =$  (default =2.5 for  $c_3$  crops and 4 for  $c_4$  crops) is a typical maximum conversion factor above ground biomass for  $c_3$  and  $c_4$  crops when the environmental conditions are all optimal.
- ' $c_3$ ' crops are: wheat, rice, cotton, barley, sunflower, oats, rye, alfalfa, pastures, sugar beet, potato and orchards. ' $c_4$ ' crops are: sorghum, millet, sugarcane and maize.

$$T_1 = 0.8 + 0.002 * T_{opt} - 0.005 * T_{opt}^2 \quad 2-28$$

$$1.185 * \frac{1}{1 + \exp(0.2 * T_{opt} - 10 - T_{mon})} * \frac{-1}{1 - \exp(-0.3 * T_{opt} - 10 - T_{mon})} \quad 2-29$$

Where:

- $\wedge$  Is the evaporative fraction of the surface energy balance
- $T_{opt} [^{\circ}C]$  is the mean air temperature during the month of maximum leaf area index or NDVI.
- $T_{mon} [^{\circ}C]$  is the mean monthly air temperature.
- The month of maximum NDVI might be different for different crops. It means both  $T_{opt}$  and  $T_{mon}$  can be calculated for different crops. Then, this output is restricted in the sense that: Not all crops can be evaluated.

pixels having mixed crops have to be classified as either being  $c_3$  or  $c_4$ , despite the error. This information is used to create a “mask” map having CLASS domain with two classes:  $c_3$  and  $c_4$ . The  $\varepsilon$  factor is evaluated per image and per crop. To do so, the periodic values of  $T_1$  and  $T_2$  per image and per crop and the calculated evaporative fraction  $\hat{\varepsilon}$  are needed.

A spatial selective operation calculates the accumulated biomass for pixel designated as  $c_3$  and  $c_4$ .

$$Bact = \text{if}(\text{mask} = "c_3", \sum \varepsilon_3(t).AcAPAR_i, \sum \varepsilon_4(t).AcAPAR_i) \quad 2-30$$

Where:

- $\varepsilon_i(t)$  Has two values per each  $AcAPAR_i$  image going into calculations. They correspond to the  $c_3$  and  $c_4$  crops for this image.
- The user enters the  $AcAPAR_i$  images. Their amount is variable, depending on availability.
- The mask is used to apply the correct  $\varepsilon$  (crop dependent) to each pixel.
- If the user selects to apply the calculations to  $c_3$  or  $c_4$  only, the mask is used again to apply the correct  $\varepsilon$  (crop dependent) to each pixel that corresponds to the selected crop only.

## 2.2 Suitability of Remote Sensing for Biomass Quantification

The tasks of remote sensing are partly geometric and partly radiometric that makes it suitable for use in gathering information of a substance with varied spatial and temporal distribution (Hartl, 1989). The geometric tasks involve: cartographic mapping of regions; thematic mapping; identification of substances, structures and regional interrelationships; recognition of processes and trends and monitoring of regional, seasonal developments. With the radiometric part, the information is contained in: spectral reflectance; scatter/backscatter varieties; emissivity as a function of materials and substances; temporal variations; structural/textural features.

Conventional ground survey techniques are subject to interpreter bias, potential logistical difficulties resulting from local topography and infrastructure and the overall expense of the survey exercise. Aerial photographic techniques can cover large areas rapidly and acquire data on a systematic basis and have a high spatial resolution (about 50 cm), but they are often prohibitively expensive. By contrast remote sensing by satellites can provide relatively homogenous sample over extensive geographical areas through a range of spatial, temporal and spectral resolutions depending on the choice of satellite medium (Hay *et al.*, 1997). Remote sensing can be used to establish the statistics of the cultivated areas, statistics on the progress of the crops, to predict the harvest, and to provide continuous and extremely prompt information on the state of growth of the main crops and the degree of damage caused by pests, disease and other agents which affect the actual yield.

The availability of up-to-date satellite geographic imagery provides valuable visual information resource that supports a wide range of Geographic Information Systems (GIS) applications. Some GIS application areas that benefit from the uses of imagery are: using current image data sources for visual reference and to update existing maps, automated land use and cover mapping, property damage assessment from recent natural hazards (floods, hurricanes, tornadoes, fires), disaster and response preparedness, agricultural crop condition and stress assessment, resource management planning and vegetation vigor assessment and biomass mapping.

GIS is a beneficial tool for biomass mapping as: it offers a ready tool for display of spatial information, provides ready algorithms for spatial analysis and modeling and gives ready tools for query and information retrieval. ArcView Image Analysis in ArcView GIS provides a direct path from ArcView GIS to ERDAS IMAGINE<sup>®</sup> for users with more complex or sophisticated geographic imaging and processing needs. Data that has been manipulated in ERDAS IMAGINE<sup>®</sup> can also be processed in ILWIS academic<sup>®</sup>.

### 2.3 Limitations of Remote Sensing

The efficacy of using remote sensing techniques is limited by the ability of the user to correlate spectral response patterns with the condition on the earth surface. The factors that limit the strength of the correlation are:

- i. Appropriateness and accuracy of image pre-processing techniques;
- ii. The adequacy of algorithm used to classify the remotely sensed images;
- iii. The suitability of ground truthing methods and
- iv. The representativeness of field sampling locations (Warren *et al.*, 1990, Matsuda, 1982).

## Chapter 3

### 3. Materials and Methods

#### 3.1 Study Area

The study was conducted in the 3,300 hectares Mumias Sugar Company Nucleus Estate situated ten kilometers to the North-East of Mumias town in Mumias-Butere district, Kenya (Figure 3-1). The town is located at Longitude 0°37'20"N, and Latitude 34° 30'E while the study area is bounded by latitude 34° 30'E to latitude 34° 50'E and Longitude 0°38'N to longitude 0°45'N (GoK, 1971).

It lies in the Lower midland sugarcane zone (LM 1 l-mi) with a very long cropping season and intermediate rains, divisible into a long cropping season followed by a medium one (Jaetzold and Schimdt, 1982). The area is very suitable for growing sugarcane. The predominant sugarcane varieties grown in the area include; CO617, CO945, KEN83-30, 8687D, 8484D, 8532D, D85-30 and D85-32.

The area receives an annual rainfall of 2000 mm in a bimodal pattern that begins in late February to July and September to early December. The maximum and minimum temperatures recorded in this area are 29.7° C and 14.4° C with a mean daily of 22.1°C. The annual evaporation is 1660 mm with an average radiation of 23.7 mJm<sup>2</sup>/day. According to Jaetzold and Schmidt (1982), a complex of soils developed on granites and quartz-feldspars gneisses which are moderately to well drained, shallow to very deep, reddish brown, to yellowish brown, friable clays over petroplinthite (orthic FERRALSOLS and partly petroferric phase with ACRISOLS) predominate the area.

The area was chosen for the study for the following reasons:

- Mumias Sugar Company produces over fifty five percent of the sugar produced in Kenya (KSB 2003, Figure 3-2) with total field coverage of over 48,000 hectares spanning over both the nucleus estate and the out grower farmer fields;

- Secondly, the company is a major user of biomass from sugarcane for cogeneration with an installed generation capacity of 5MW (Ministry of Energy, 1992). This makes mapping of sugarcane biomass in the area significant;

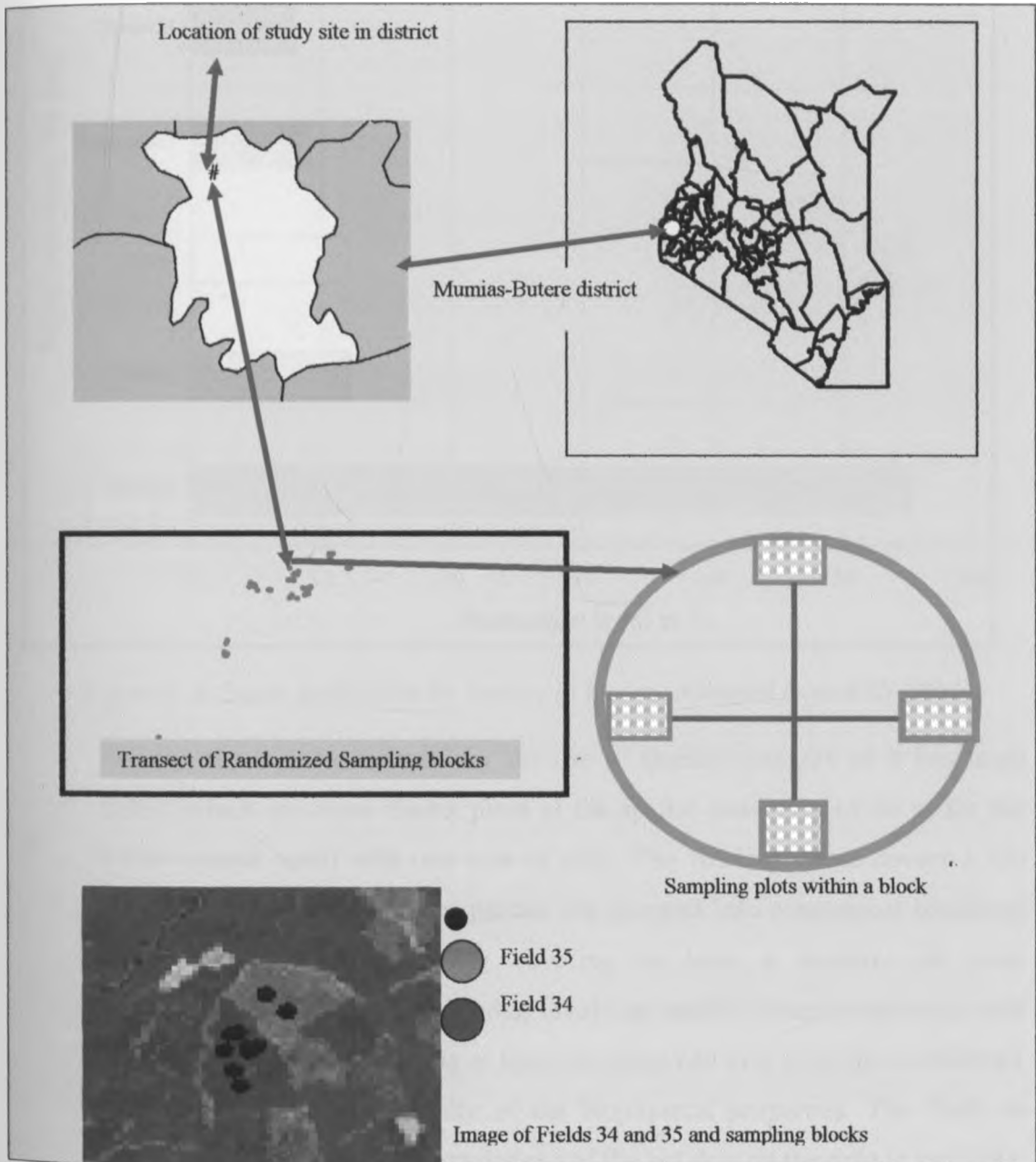


Figure 3-1: Study area

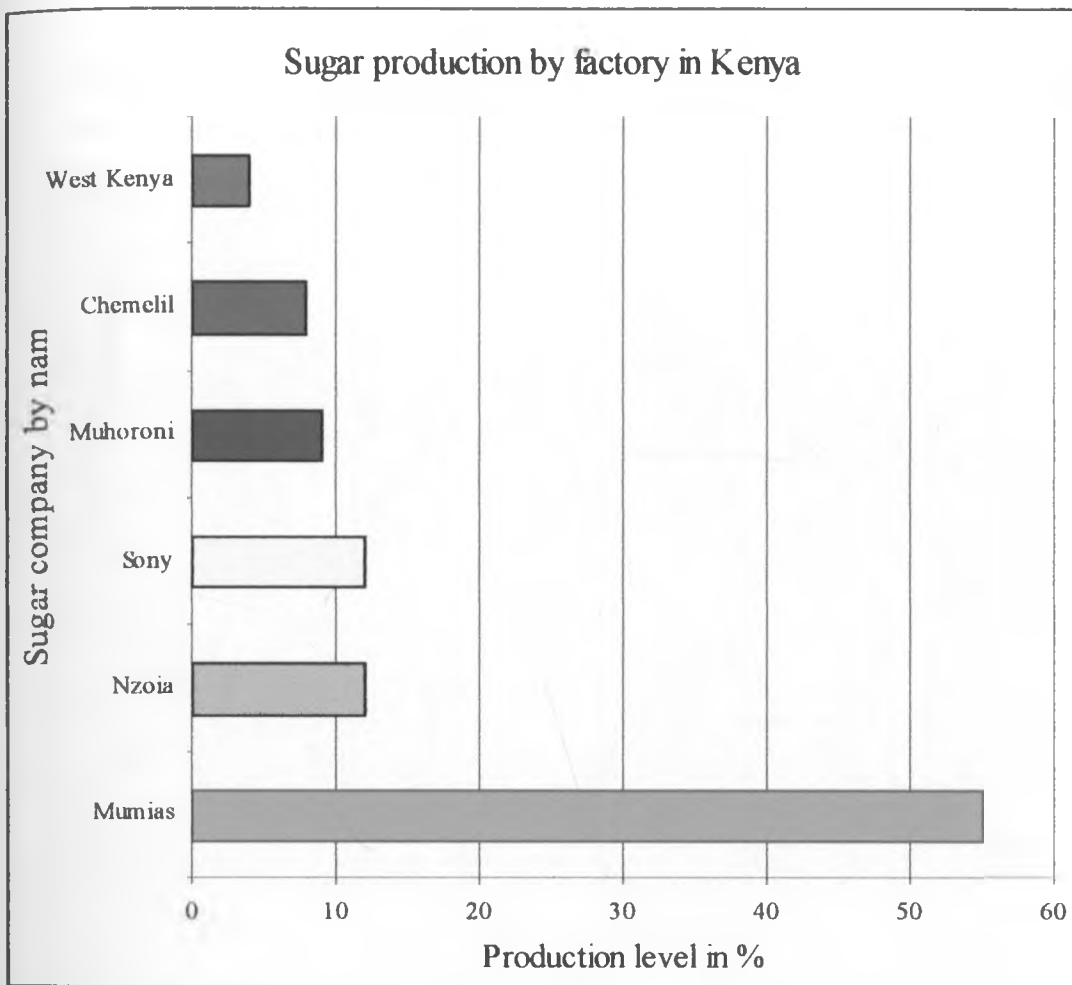


Figure 3-2: Sugar production by factory in Kenya. *Adapted from KSB 2003*

- Thirdly, the area lends itself for the use of satellite imagery as it has large fields (which are more than a pixel at the spatial resolution of 30 m for the ETM+ sensor used) with one type of crop. The Nucleus estate covers 3,300 hectares while the out grower parcels are grouped into economical blocks of sugarcane fields (Figure 3-3) covering at least 6 hectares of land. Recommended field sizes for studies involving satellite imagery are areas with features to be studied covering at least one pixel (30 m x 30 m for LANDSAT ETM+) to ensure homogeneity of the biophysical properties. The fields in Mumias have well marked boundaries and the bio data on the crop is available. The Nucleus Estate also provides an area with a uniform management system for the crop reducing the effects of errors due to different crop management styles;

- Fourthly, the area straddles a sugarcane ecological zone, making it an ideal site for a study on sugarcane (Jaetzold and Schmidt, 1982).



Figure 3-3: Sugarcane fields in Mumias Nucleus Estate (*mature, growing and harvested*)

- Fifthly, the area lies within the densely populated areas of the Lake Victoria basin where poverty lurks on humanity (Omuto 2002, ICRAF, 2000, Van der Kwast, 2002) and where scientific interventions are being sought to increase



understanding on factors that influence production of agricultural crops. In this area, the net earnings from sugarcane have not kept pace with inflation and economic growth and there is need to improve production while lowering costs (KSB, 2002). A study that will give a rapid methodology for predictive spatial and temporary mapping of sugarcane biomass is a positive step towards lowering the costs.

- Sixthly, the area is a locus for various studies by many international and research organizations making acquisition of secondary data a feasible endeavor.
- Lastly, the crop selected produced 2.3 million tones of the total 3.6 million tones of biomass available for industrial use in the country (Hankins, 1989) signifying its economic importance. Further, the sugarcane variety studied is the most predominant in the area (KSB, 2003).

The study concentrated on the upland section of the nucleus estate to minimize the effect of topography and soil variability. The management practices of the major study plots included primary tillage using the disc plow; secondary tillage through disking; hand planting and top dressing with Di-Ammonium Phosphate (DAP) and Calcium Ammonium Phosphate (CAN) at 5 months of age. All the fields were subjected to the same crop husbandry practices and therefore it was assumed that there was no significant variation attributed to fertilizer application, weeding, planting, plowing and harrowing across the study fields.

### **3.2 Study Protocol**

The methods used in this study can be summarized as:

- i. Selection and measurement of field data;
- ii. Selection and processing of spectral data;
- iii. Statistical analysis to determine relationships between the spectral and field data; and
- iv. Mapping (Spatial Quantification) of sugarcane biomass.

### 3.2.1 Selection, sampling design and measurement of field data

Stratified sampling is normally used if the sampled area (or volume) is heterogeneous. If the patch size is much larger than inter-sample distance, then kriging can be used instead of stratified sampling. The main advantages of kriging are that: it can handle spatial autocorrelation, it is not sensitive to preferential sampling in specific areas and that it estimates both local population densities and block averages. It can replace stratified sampling if the size of aggregations is larger than the inter-sample distance. The methods used with kriging allow it to have an advantage over other estimation procedures in that the estimated values have a minimum quantifiable error associated with them. The stratified approach suited the Mumias sugar fields since clipping was done within 6-9 ha fields.

In stratified sampling program, the area (volume) is subdivided into two or more portions, which are sampled separately. The Randomized Block Designs is a research design, which is equivalent to stratified sampling. The randomized block designs are constructed to reduce noise or variance in the data. This is achieved by dividing the sample into relatively homogeneous subgroups or homogeneous subgroup to ensure that the variability within each block is less than the variability of the entire sample. Thus each estimate of the treatment effect within a block is more efficient than estimates across the entire sample (Trochim, 2002).

The study fields were stratified on the basis of age to obtain more accurate estimates with less effort (RCSSMRS, 1988). Fields from the upper zone of the Mumias Nucleus Estate with nearly similar soil types and drainage patterns were used. This was intended to minimize variability that would otherwise be caused by varied topography, soil and drainage. The approach taken was meant to minimize the number of replications necessary to a manageable limit as dictated by the available time, capital and human resources.

Field data was collected for sugarcane at the fifteenth (Field 34) and sixteenth (Field 35) months of growth. The ages were selected to coincide with the critical sugarcane

phenological stages of maturity and senescence. For each field and date, plant samples were taken and their structural parameters measured. Four different varieties were to be studied. However, due to time and resource constraints, biomass was collected at only 36 of the proposed 82 sites in the two Fields (34 and 35).

The two sugarcane fields studied (Figure 3-1) were of sizes 9 ha and 15 ha respectively. These were large enough to allow for classification and spectral analysis to be done pixel by pixel (each pixel covers 900 m<sup>2</sup> for LANDSAT ETM+). For each field of study, clipping was done in blocks randomly placed on a W shaped transect (

Figure 3-4). These blocks were located in the field as shown in Figure 3-1. Field 34 had 5 blocks while Field 35 had 4 blocks. Each block was 20 m by 20 m in size to ensure that it fitted within a single LANDSAT ETM+ pixel (

Figure 1-2 and Figure 3-1). The nine randomized sampling blocks were used as they meet the minimum number recommended by Congalton ((1986) i.e. number of sites should be greater than  $n+1$ , where  $n$  is the number of bands of the imagery used. In this case,  $n$  was 6).

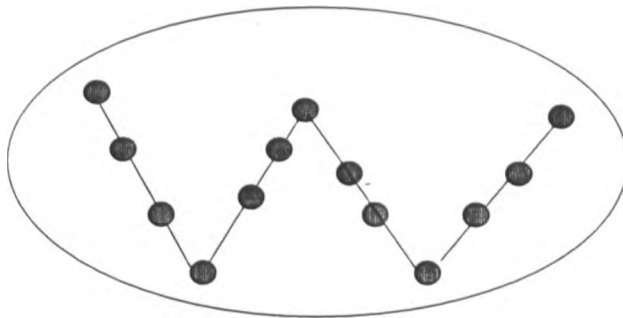


Figure 3-4: W-shaped sampling transect

Given an error of less than one pixel in positional accuracy for the satellite imagery, an exclusion buffer of 30 m, or one pixel, was applied when extracting the at satellite reflectance pixel values for each field to reduce edge effects.

From each block, four 1.5m x 1.5m sampling plots were clipped for biomass study along a  $\dagger$  shaped transect (

Figure 1-2 and Figure 3-1). Detailed sampling was done for cane to provide biophysical (cover and spatial biomass variability, average plant height, the weight of wet and dry biomass and the leaf water content and leaf specific weight) in-situ data for correlation with spectral data. Besides, cultivation data and selected meteorological data were recorded. Accurate field boundary information for the nucleus estate farms was obtained for the 2002-2003 season. Expected sugarcane yield, fertilization, plowing details, seeding date and density, variety, and harvest date information was obtained for these fields.

Hand held Global Positioning System (GPS) reading was obtained at the center of each sampling plot using Garmin® 12 XL GPS unit (Figure 3-5) to establish plot locations in Kenyan Universal Transverse Mercator (UTM) coordinates based on WGS 84 datum. The geographic location of each sample site was recorded within a positional accuracy of 0.15 pixels or 4.5 m. The choice of GPS equipment to use was considered on the basis of its capability to provide coordinates in the UTM system automatically as opposed to the more tedious conventional systems which would have required the measurement of angles and distances to obtain the coordinates. Additionally, the GPS handset is a satellite based measuring system which does not depend on weather for its operation. Garmin® 12 XL handsets are known to give readings with accuracies of 1 to 3 m making them suitable for ecological sampling in plots that are homogeneous within 2 to 5 meter spans.

### **3.2.2 Selection and processing of spectral data**

One cloud-free, LANDSAT ETM+ image of February 22, 2003 for path 169, row 70 covering the study area was obtained from World Agro forestry Center (ICRAF) Kenya. The intention was to select a satellite image covering a period with low cloud cover and minimal constraints in conducting fieldwork. Field data was gleaned in the months of January, February and March 2003, but only the February 2003 image was acquired. The January and March images were not available at the World

Agroforestry Center and therefore could not be used with their respective ground truthed data. However, the March ground truthed data were used with the February acquired image to test the strength of the correlations for data gleaned beyond the satellite overpass period.

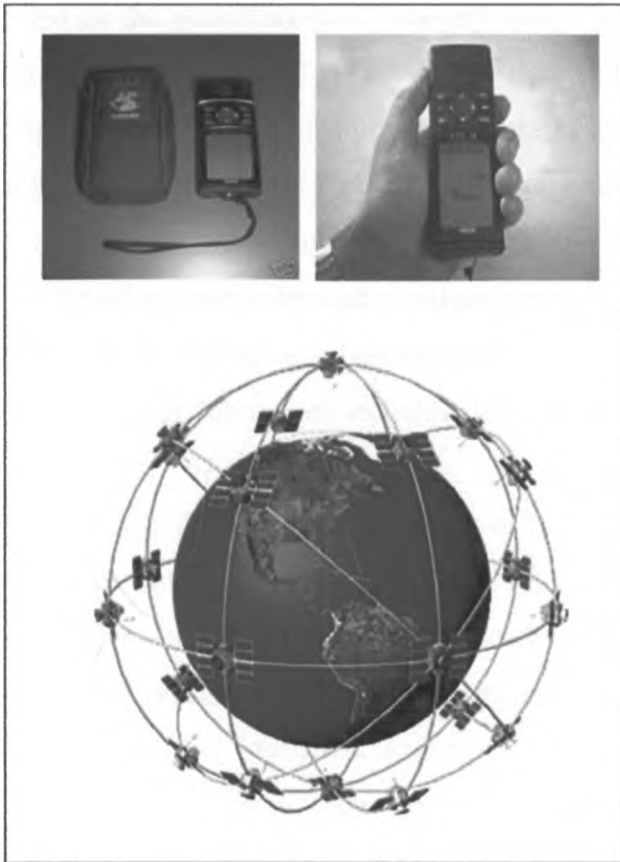


Figure 3-5: The Global Positioning System handsets and a network of satellites

The LANDSAT ETM+ image for February 2003 was geo-referenced and corrected for geometric error by first transforming the February image to the UTM projection in zone 36. The geometric transformation equation was computed using six ground control points that produced a final root mean square (RMS) error of 0.15 pixels. This error was less than 0.35 pixels recommended for ecological studies. The geo-referencing was done using ground control points obtained from two 1:50 000 digital Toposheets (GoK, 1971) augmented with four ground acquired GPS coordinates. The GPS points used were marked at; the road junctions at Mayoni market and Mumias town center, the bridge crossings on river Nzoia at Mayoni on Bungoma-Kakamega

road and at Yala on Kisumu-Busia road. Georeferencing was done using ERDAS® 8.4 software. This georeferenced image was then imported into ILWIS® for spectral analysis.

Field site locations were digitized as a vector file using the GPS readings in ArcView 3.2a and then imported on the geo-referenced February image in ILWIS. The ETM+ bands used were bands 1 (blue-green), 2 (green), 3 (red), 4 (NIR), 5 (MIR), and 6 (MIR) (Table 3-1). Spectral data was obtained from the dataset for 36 sample plots, 20 of which were from Field 34 and 16 from Field 35. The image was then examined for sensor errors (i.e., striping, banding, and line dropout) and the digital numbers converted to reflectance and contrast enhanced to values between zero and 100 to give a physical representation. The reflectance values were adjusted for atmospheric scatter using the Improved Dark Object Subtraction technique described by Chavez (1996).

The spectral values for each plot were interpolated using a nearest neighbor re-sampling approach and the data were output to a 30 x 30 m pixel size. During the geometric correction process, the image was clipped to the Mumias Nucleus Estate study field boundaries. The radiance values for each pixel in the sample sites were extracted and stored in excel for statistical analysis (i.e. distribution, mean and standard deviation). Field information on crop variety, yield, and cumulative monthly age at the image acquisition dates were stored in excel for analysis.

Spectral vegetation indices, ratios and biophysical properties were computed using scripts written and run in ILWIS 3.2 academic<sup>®</sup> (2005) to solve equations for estimating biophysical and agro-ecological properties of sugarcane and to produce map formats of the properties from the February 2003 LANDSAT ETM+ satellite image. The choice of the software to use was mainly governed by its friendliness and capabilities to use written scripts to extract biophysical properties of vegetation from satellite imagery.

The indices, ratios and biophysical parameters extracted from the image were: Normalized Difference Vegetation Index (NDVI) and Soil Adjusted Vegetative Index (SAVI), (R/NIR), NIR/MIR and R/MIR (Crist *et al.*, 1986). Others were LAI, fPAR, fcover, and accumulated biomass. The computations were done using a script written and run in ILWIS ® software (appendix A.1). Simple linear regression models to predict crop biophysical parameters based on these vegetation indices (VI) were developed. Vegetation index values, resulting from all possible combinations of the selected wavebands, were used to develop prediction models. The indices were used to highlight the vegetation component in a soil background and, due to normalization, to minimize the effects of illumination and other measurement errors. This allowed for spatial comparison of the same crop.

Table 3-1: The spectral bands: LANDSATETM+ satellite (Irish, 2000 and Noonan, 1999).

ETM+ Spectral band widths ( $\mu\text{m}$ )	BAND 1	BAND 2	BAND 3	BAND 4	BAND 5	BAND 6
	0.45-0.52	0.53-0.61	0.63-0.69	0.78-0.90	1.55-1.75	10.4-12.5
Ground resolution (m)	30	30	30	30	30	60
Electromagnetic region	Visible Blue	Visible green	Visible Red	Near infrared	Middle infrared	Thermal
Generalized application details	Coastal water mapping, differentiation of vegetation from soils	Assessment of vegetation vigor	Chlorophyll absorption for vegetation differentiation	Biomass surveys and delineation of water bodies	Vegetation and soil moisture measurements	Hydrothermal mapping
Used in analysis	Yes	Yes	Yes	Yes	Yes	Yes



### 3.2.3 Statistical (Geostatistical) analysis

Geostatistics is a collection of statistical methods which were traditionally used in geo-sciences. These methods describe spatial autocorrelation among sample data and use it in various types of spatial models. Geostatistical methods have recently been adopted in ecology and appear to be very useful in this new area and have changed the entire methodology of sampling. Traditional sampling methods are not compliant with auto-correlated data and therefore, the main purpose of sampling plans is to avoid spatial correlations. In geostatistics there is no need in avoiding autocorrelations and sampling becomes less restrictive.

Also, geostatistics changes the emphasis from estimation of averages to mapping of spatially-distributed populations. Spatial autocorrelation can be analyzed using correlograms, covariance functions and variograms (semivariograms). In brief, geostatistical analysis usually has the following steps:

- Estimation of correlogram
- Estimation of parameters of the correlogram model
- Estimation of the surface (=map) using point kriging, or
- Estimation of mean values using block kriging

Detailed description of most geostatistical methods can be found in Isaaks *et al.* (1989).

The General Linear Model (GLM) underlies most of the statistical analyses that are used in applied and social research. It is the foundation for the t-test, Analysis of Variance (ANOVA), Analysis of Covariance (ANCOVA), regression analysis, and many of the multivariate methods including factor analysis, cluster analysis, correlation, Principal Component Analysis and Canonical or Discriminant Analysis.

Pearson's Correlation analysis was used to determine the linear relationships between the estimated biophysical parameters and measured biomass. Pearson's Correlation Coefficient measures the strength of the linear relationship between two variables. It assumes that both the variables (often called X and Y) are interval/ratio and

approximately normally distributed, and that their joint distribution is bivariate normal. Pearson's Correlation Coefficient is usually signified by  $r$  (rho), and can take on the values from -1.0 to 1.0. Where -1.0 is a perfect negative (inverse) correlation, 0.0 is no correlation, and 1.0 is a perfect positive correlation. The related statistics to Pearson's Correlation Coefficient is  $R^2$  (called the coefficient of determination or  $r$  squared) which is normally interpreted as the proportion of variance in  $Y$  that is contained in  $X$ . The statistical significance of  $r$  is tested using a t-test. The hypotheses for this test are:

$$H_0: \rho = 0$$

$$H_a: \rho \neq 0$$

A low p-value for this test (less than 0.05 for example) means that there is evidence to reject the null hypothesis in favor of the alternative hypothesis, or that there is a statistically significant relationship between the two variables.

The correlation coefficient  $r$  (also called Pearson's product moment correlation) is calculated by

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

3-1

Assumptions:

- linear relationship between  $x$  and  $y$
- continuous random variables
- both variables must be normally distributed
- $x$  and  $y$  must be independent of each other

Equation 3-1 can be replaced by an equivalent formula (Equation 3-2) which avoids the use of the means and is therefore much faster to calculate.

$$r = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2} \sqrt{\sum_{i=1}^n y_i^2 - \frac{1}{n} \left( \sum_{i=1}^n y_i \right)^2}} \quad 3-2$$

The correlation coefficient stands in close relationship to linear regression. The square of  $r$  is called the goodness of fit and denotes the portion of total variance explained by the regression model. Correlation coefficients were listed using a correlation matrix table. A mean transformation was used to normalize the data distribution of the biomass and checked for estimation bias using Jack-Knife re-sampling technique. The correlation and regression analyses were performed using the February ETM+ imagery. All relationships between biophysical properties and measured biomass were made using the February ETM+ imagery. The biophysical measurements were however made in January, February and March 2003.

According to Congalton *et al.* (1998), at least 50 samples for each category should be used to evaluate classification accuracy. Due to the small sample size of 36 associated with this biomass study, the accuracy of the function was tested using a Jack-Knife Cross Validation (JCV) approach. Performance of the developed models was evaluated by comparing the observed and predicted values. The first set (the estimated calibration data) was used to develop the prediction models, while the second data set (the measured validation data) was used to test the performance of the models. Jack-Knife is used for bias removal. The Root Mean Square (RMS) for an estimator is given as equal to the square of the bias plus the variance of the estimator. If the bias is much higher than variance then under some circumstances Jack-Knife could be used. The first application of Jack-Knife can reduce bias without changing variance of the estimator. But its second and higher order application can in general increase the variance of the estimator.

To carry out regression analyses to generate models, the spectral responses of the crops (i.e., the reflectance values recorded in various wavebands) were considered as independent parameters, and the biophysical parameters were the dependent

parameters. Analysis of variance (ANOVA) was done on both the calibration and validation data while the data was tested for accuracy using the student t-test and the F critical test at 0.05 removal.

The sum of squared error (SSE) and the average relative percent error (ARPE) values were calculated for both calibration and validation data sets, while the more stringent coefficient of efficiency (or Nash-Sutcliffe coefficient),  $C_{eff}$  (James *et al.*, 1982), was calculated for the validation data only:

$$SSE = \frac{\sum_{i=1}^{i=n} (S_i - O_i)^2}{n}$$

3-3

Where:

$O_i$  = individual observed value

$S_i$  = individual simulated value

SSE = Sum of Squared Error

$n$  = number of paired observed-simulated values

The SSE is an indicator of the quantitative dispersion between the observed and simulated values, while the ARPE expresses the error and the sign of the error indicates whether the model over- or under-predicted the values.

The  $C_{eff}$  evaluates the error relative to the natural variation in the observed values. A  $C_{eff}$  of 1.0 represents a perfect prediction, a value of 0 (zero) represents a prediction no better than simply using the observed mean as a prediction, and increasingly negative values indicate increasingly poorer predictions.

$$ARPE = \frac{\sum_{i=1}^{i=n} (S_i - O_i)}{\sum_{i=1}^{i=n} O_i}$$

3-4

$$C_{eff} = \frac{\sum_{i=1}^{i=n} (O_i - \bar{O})^2 - \sum_{i=1}^{i=n} (S_i - O_i)^2}{\sum_{i=1}^{i=n} (O_i - \bar{O})^2}$$

3-5

Where:

$O_i$  = individual observed value

$S_i$  = individual simulated value

$O$  = mean observed value

$n$  = number of paired observed-simulated values.

### 3.2.4 Spatial quantification and mapping of sugarcane biomass

The simple linear regression models performed returned LAI and fPAR as the best predictors of biomass using satellite data. The fPAR satellite scene values were used to estimate biomass and to derive a linear relationship for estimating accumulated biomass based on equation 2-26

The equation was then solved using a biophysical script written and run in ILWIS software to estimate accumulated biomass based on filed information acquired from the satellite image for the study area. The map output of biomass in dry weight was based on quantified biomass yield of sugarcane at the age of 16 months using information derived from the ILWIS output. Mapping was done to provide a visual picture of the integration of interaction of factors affecting biomass yield distribution.

## 4. Results and Discussions

### 4.1 Field and Laboratory Data

The measured field data for 36 sampling plots are given in table 4-1. Four blocks with four sampling plots of 2.25 m<sup>2</sup> areas each were gleaned from Field 35 in February 2003 while five similar blocks were sampled from Field 34 in March 2003. A February 22, 2003 LANDSAT ETM+ image was acquired and used for analysis. The March 2003 image was not available at World Agroforestry center at the time of analysis and was therefore not acquired.

Table 4-1: Measured Biomass data for Fields 34 and 35 at the sampling plots

Field 34					
Block No.	Plot No.	Biomass t/ha	Block No.	Plot No.	Biomass t/ha
1	41	196	3	61	133
	42	213		62	138
	43	178		63	111
	44	173		64	111
2	51	160	4	71	267
	52	156		72	111
	53	178		73	133
	54	93		74	169
5	81	164	5	83	164
	82	124		84	182
Field 35					
Block No.	Plot No	Biomass t/ha	Block	Plot No	Biomass t/ha
1	11	120	3	31	196
	12	116		32	148
	13	133		33	164
	14	147		34	203
2	22	130	4	41	200
	23	120		42	187
	24	116		43	200
				44	240

### 4.2 Crop reflectance data extraction

Spectral values extracted from the acquired LANDSAT ETM+ February 22, 2003 image were transformed into Vegetative Indices and band ratios (NDVI, SAVI, R/NIR, R/MIR and NIR/MIR). The spectral values were also used to derive

biophysical parameters (f-PAR, f-Cover and LAI, and Accumulated Biomass, Table 4-2 and Table 4-3) at the sampling plots using the script written and run in ILWIS (Appendices). The extracted spectral values and the transformed indices and biophysical parameters are listed in Table 4-2 and Table 4-3 for Fields 34 and 35 respectively. Field 35 was sampled when at senescence stage while Field 34 was sampled at an active growing stage. Field 34 had higher chlorophyll content than Field 35 and therefore showed the brighter red color than that of Field 35 (Figure 4-1) using a Red-Green-Blue (RGB) color projection.

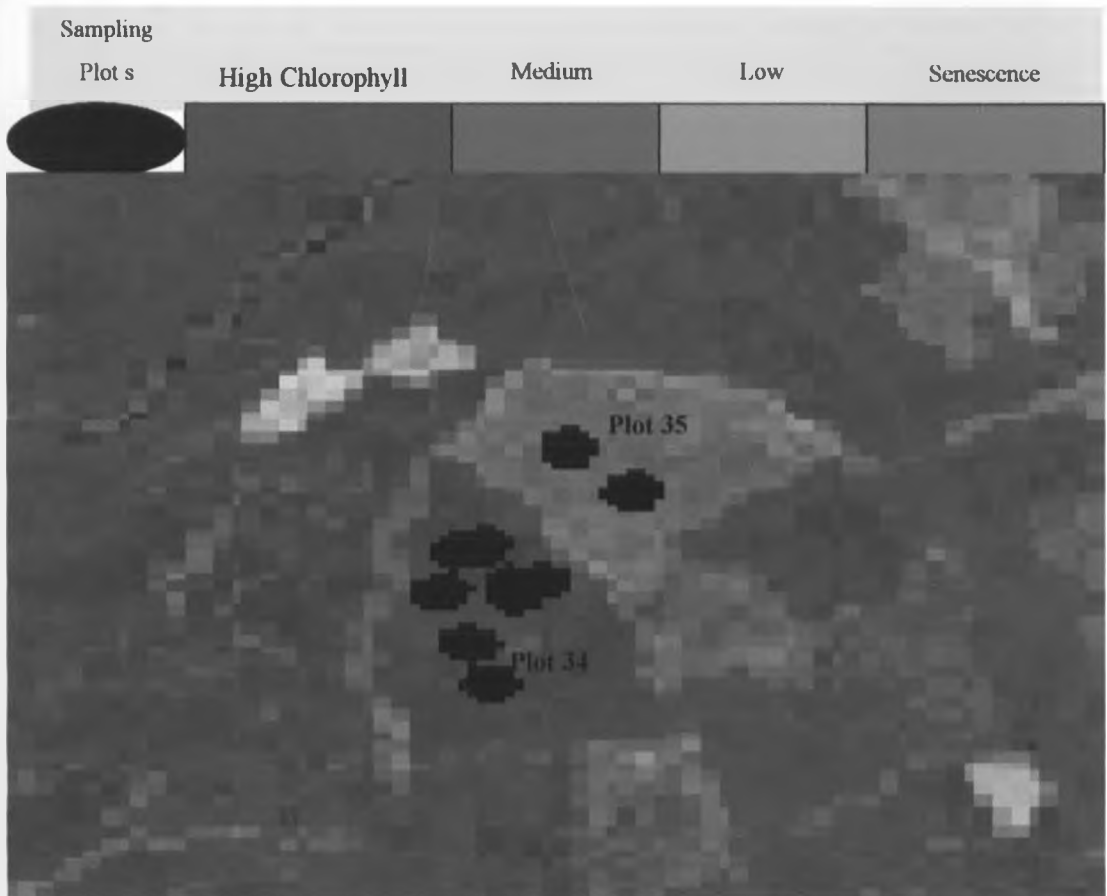


Figure 4-1: Geo-referenced image clipped to show plots 34 and 35

Table 4-2: Spectral analysis and biophysical estimation for Field 34

Plot	Estimated biophysical Parameters											Biomass t/ha	
	fPAR	B1	B2	B3	B4	B5	B6	fCover	NDVI	SAVI	LAI	Measured	Estimated
B3411	0.53	52.21	43.93	26.70	70.51	8.21	4.50	0.38	0.59	0.35	0.13	196	134.026
B3412	0.54	54.54	43.93	27.95	74.38	9.60	4.87	0.39	0.6	0.36	0.13	213	133.499
B3413	0.56	49.87	42.33	24.84	69.54	7.58	4.39	0.41	0.61	0.35	0.13	178	130.337
B3414	0.53	52.21	43.93	26.70	70.51	8.21	4.50	0.38	0.59	0.35	0.13	173	134.026
B3421	0.55	51.43	43.13	24.22	67.60	7.08	4.50	0.4	0.61	0.34	0.12	160	131.039
B3422	0.49	54.54	44.72	31.68	74.38	10.74	5.02	0.34	0.55	0.34	0.13	156	140.701
B3423	0.49	54.54	44.72	31.68	74.38	10.74	5.02	0.34	0.55	0.34	0.13	178	140.701
B3424	0.49	54.54	44.72	31.68	74.38	10.74	5.02	0.34	0.55	0.34	0.13	93	140.701
B3431	0.54	50.65	43.13	26.70	71.47	8.97	4.80	0.39	0.6	0.35	0.13	133	133.323
B3432	0.52	49.87	43.93	26.08	65.66	8.09	4.43	0.37	0.58	0.33	0.12	138	136.309
B3433	0.52	49.87	43.93	26.08	65.66	8.09	4.43	0.37	0.58	0.33	0.12	111	136.309
B3434	0.54	50.65	43.13	26.70	71.47	8.97	4.80	0.39	0.6	0.35	0.13	111	133.323
B3441	0.54	52.21	42.33	26.08	68.57	7.96	4.54	0.39	0.59	0.34	0.12	267	133.674
B3442	0.53	50.65	41.53	26.70	70.51	8.97	4.84	0.38	0.59	0.35	0.13	111	134.026
B3443	0.53	50.65	41.53	26.70	70.51	8.97	4.84	0.38	0.59	0.35	0.13	133	134.026
B3444	0.53	50.65	41.53	26.70	70.51	8.97	4.84	0.38	0.59	0.35	0.13	169	134.026
B3451	0.50	51.43	39.13	26.08	63.72	8.85	4.54	0.36	0.57	0.31	0.11	164	138.242
B3452	0.58	52.98	41.53	24.22	72.44	8.34	4.32	0.43	0.63	0.37	0.13	124	127.175
B3453	0.49	51.43	41.53	27.95	66.63	8.72	4.80	0.34	0.56	0.32	0.12	164	139.998
B3454	0.50	51.43	39.13	26.08	63.72	8.85	4.54	0.36	0.57	0.31	0.11	182	138.242



Table 4-3: Spectral analysis and biophysical estimation for Field 35

Plot	Extracted spectral and biophysical parameters from LANDSAT image										Measured In-situ Values		
	B1	B2	B3	B4	B5	B6	f-Cover	NDVI	SAVI	Estimated Biomass t/ha		Measured Biomass t/ha	
										All plots	Normalized (mean)	All plots	Normalized (mean)
B3511	49.1	40.7	25.5	69.5	7.3	4.5	0.4	0.6	0.3	132	134	120	118.
B3512	49.9	42.3	26.1	66.6	8.0	4.6	0.4	0.6	0.3	135		116	
B3513	49.9	42.3	26.1	66.6	8.0	4.6	0.4	0.6	0.3	135	134	133	140
B3514	49.1	40.7	25.5	69.5	7.3	4.5	0.4	0.6	0.3	132		147	
B3521	49.1	39.9	24.2	68.6	8.0	4.5	0.4	0.6	0.4	130	130	140	135
B3523	49.1	39.9	24.2	68.6	8.0	4.5	0.4	0.6	0.4	130		130	
B3524	49.1	39.9	24.2	68.6	8.0	4.5	0.4	0.6	0.4	130	130	120	118
B3531	49.1	39.9	24.2	68.6	8.0	4.5	0.4	0.6	0.4	130	130	116	
B3532	51.4	43.9	36.7	47.2	11.9	5.9	0.1	0.3	0.2	183	183	196	177
B3533	51.4	43.9	36.7	47.2	11.9	5.9	0.1	0.3	0.2	183		148	
B3534	51.4	43.9	36.7	47.2	11.9	5.9	0.1	0.3	0.2	183	183	164	183.5
B3541	51.4	43.9	36.7	47.2	11.9	5.9	0.1	0.3	0.2	183		203	
B3542	53.8	43.1	37.3	41.4	11.5	6.0	0.0	0.2	0.1	195	193	200	193.5
B3543	52.2	43.1	35.4	41.4	11.9	6.1	0.1	0.3	0.1	191		187	
B3544	53.8	43.1	37.3	41.4	11.5	6.0	0.0	0.2	0.1	195	195	200	200.

The spectral response for the sugarcane canopy is illustrated in figures 4-2 and 4-3. The figures show the normal trend of high absorption at the visible and middle infrared band levels, and high reflectance at the near-infrared band level. This conforms to what is espoused by Guo, *et al.* 1996 as depicted in figure 2-1. For the Mumias Sugarcane field, when incident solar radiation interacts with green leaves the level of absorption by chlorophyll and carotenoid pigments is approximately 50% to 80%, of the visible light in the 0.3-0.7 $\mu$ m region. In the 0.74–1.1 $\mu$ m region, there is high reflectance and transmittance of the near-infrared radiation from the leaves. In this region, the green sugarcane vegetation canopy reflected 70%- 75% of the incident near-infrared energy while the senescent canopy reflected 47- 70% of the energy. In the 1.3-12.5 $\mu$ m regions, there was absorption due to the presence of liquid water (moisture content of 66%), which caused the general decrease of reflectance for mid-infrared bands for both the green and senescent canopies (Figures 4-2 and 4-3). The first phenomenon was due to the presence of active chloroplasts while the second was explained by the scattering of radiation caused by the leaf structure. Most of the chloroplasts are present in the upper layers of leaves (In the palisade layer just below the transparent cuticle and upper epidermis). Thus visible light was either absorbed (80% for B3 (red) and 60% for B1 (blue light), and) or reflected for B2 (green light, 50%), Figure 4-2, from these layers in the leafy sugarcane canopy.

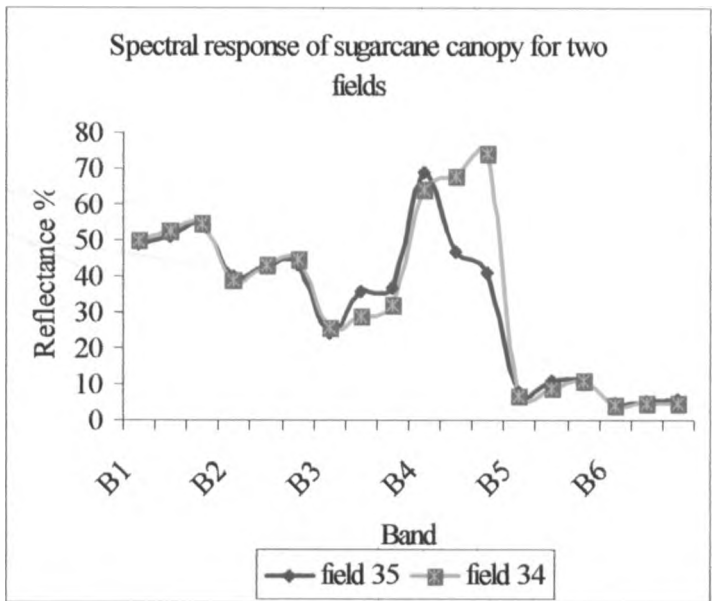
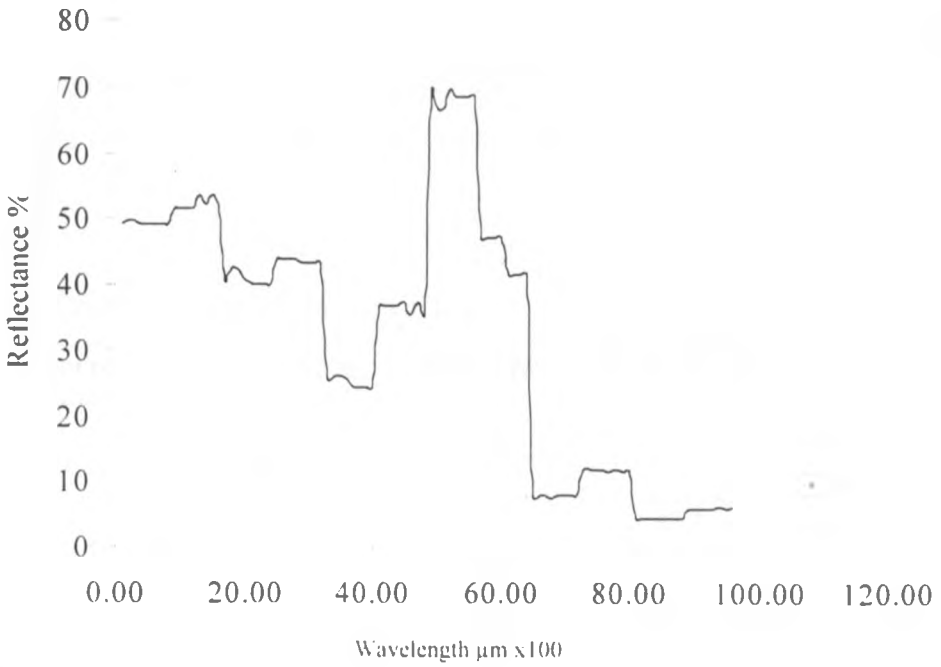


Figure 4-2: Spectral signatures of sugarcane canopy in Fields 34 and 35

### Spectral Signature



### Spectral response of sugarcane canopy for field 35

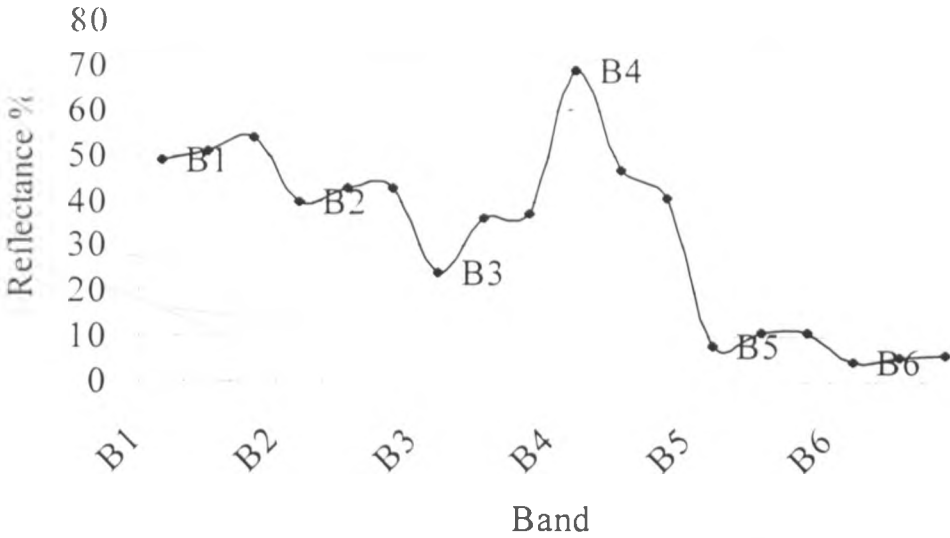


Figure 4-3: Spectral response of Mumias sugarcane canopy at senescence (Field 35)

### 4.3 Quantifying Sugarcane Biomass using Regression Analysis

Correlation values showing the strength of the relationships between selected biophysical and measured biomass are listed in Table 4-4 and illustrated in Figure 4-4 to Figure 4-12. Figure 4-4 and Figure 4-5 illustrates the relationship between measured biomass and the spectral values and shows prediction levels of biomass for each spectral band when the sites were considered bi-pixel wise and pixel by pixel for Field 35.

Table 4-4: Coefficient of correlation values ( $r^2$ )

<b>Correlation values between Measured biomass and band values</b>									
Variable »	n	Measured Biomass	B1	B2	B3	B4	B5	B6	
35 bi pixel	8	0.92	0.91	0.70	0.89	0.94	0.88	0.92	
35 pixel wise	4	0.98	0.93	0.71	0.94	0.99	0.97	0.99	
35 All plots	16	0.81	0.76	0.55	0.78	0.82	0.75	0.70	
34 bi pixel	10	0.0138	0.01	0.03	0.023	0.040	0.006	0.012	
34 All plots	20	0.00082	0.05	0.01	0.01	0.01	0.004	0.003	
Both Fields	36	0.2	0.21	0.053	0.176	0.176	0.13	0.16	
<b>Correlation values between measured biomass and biophysical parameters</b>									
Variable	n	NDVI	SAVI	LAI	fCover	fPAR	R/MIR	R/NIR	NIR/MIR
35 bi pixel	8	0.931	0.932	0.94	0.93	0.93	0.67	0.94	0.93
35 pixel x pixel	4	0.985	0.987	0.99	0.983	0.985	0.69	0.987	0.992
Field 35 All	16	0.82	0.818	0.818	0.816	0.82	0.554	0.825	0.808
34 Pixel by pixel	10	0.133	0.006	0.06	0.16	0.014	0.0548	0.0165	0.0386
34 All	20	0.003	0.001	0.01	0.004	0.003	0.002	0.0001	0.004
Both Fields	36	0.196	0.191	0.183	0.195	0.21	0.123	0.197	0.177

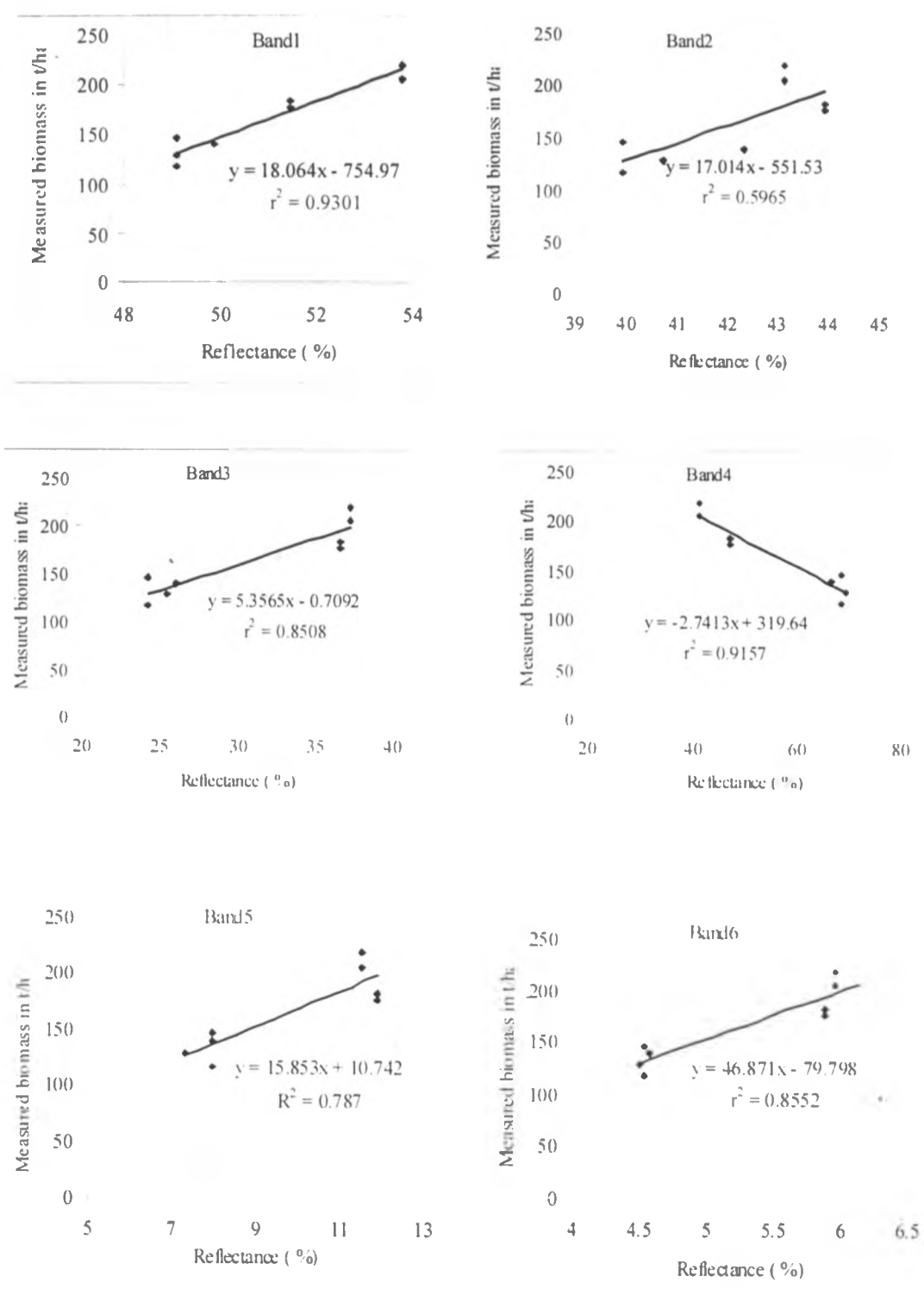


Figure 4-4: Correlation coefficients for band values (bi-pixel mean values)

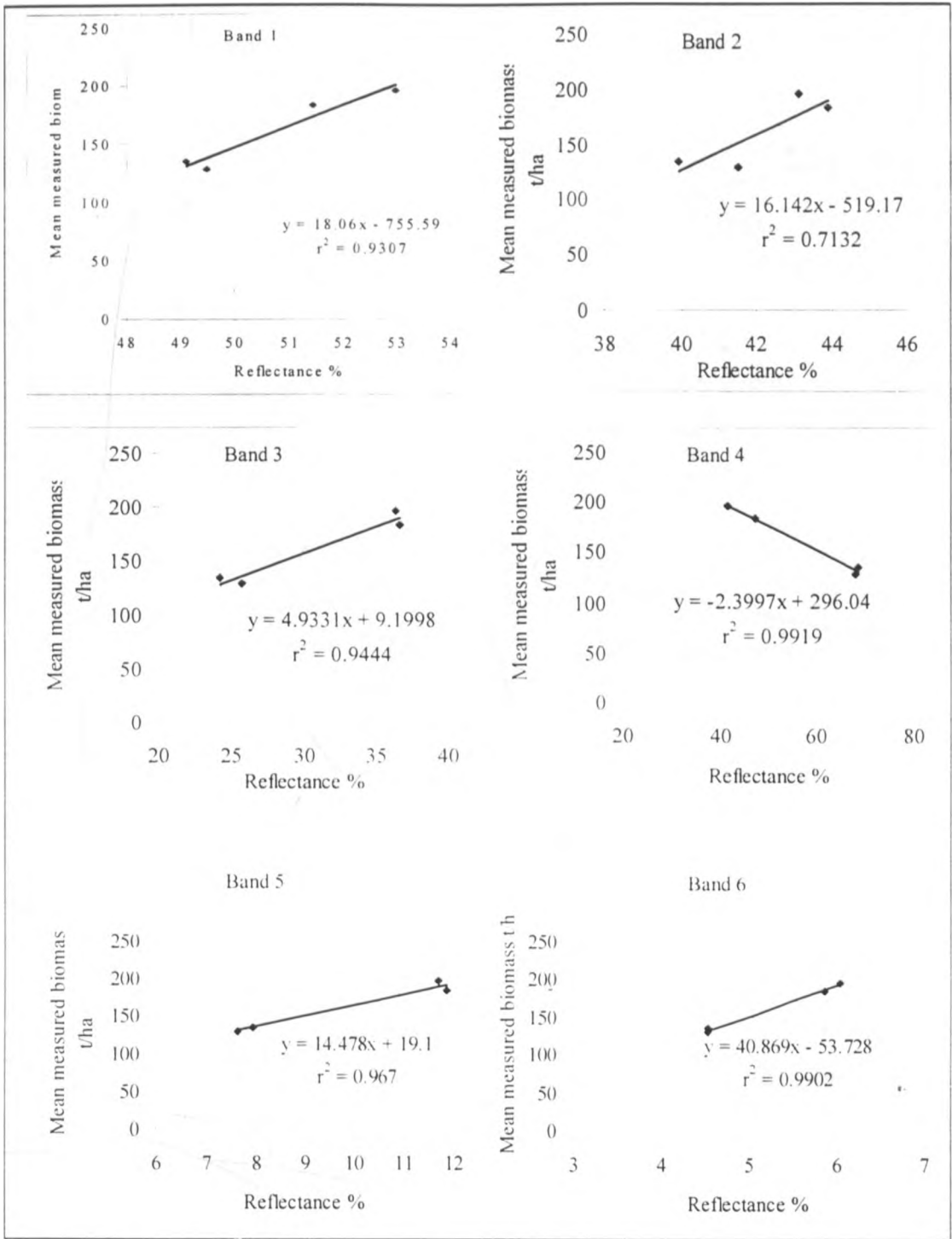


Figure 4-5: Correlation coefficients for band values (pixel by pixel mean values)

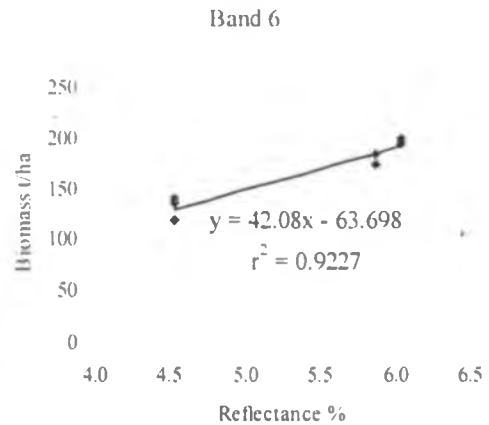
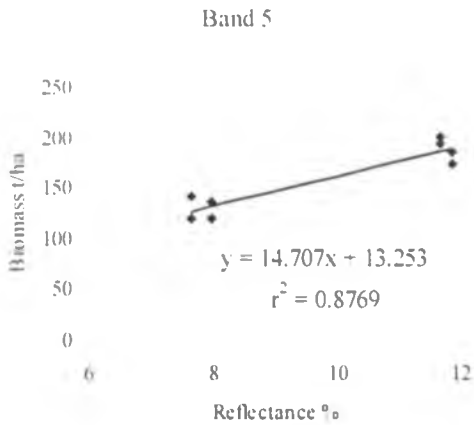
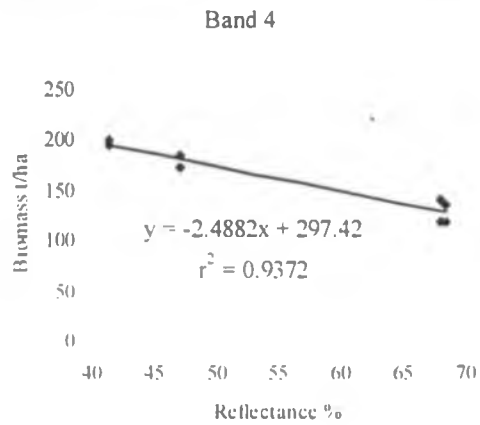
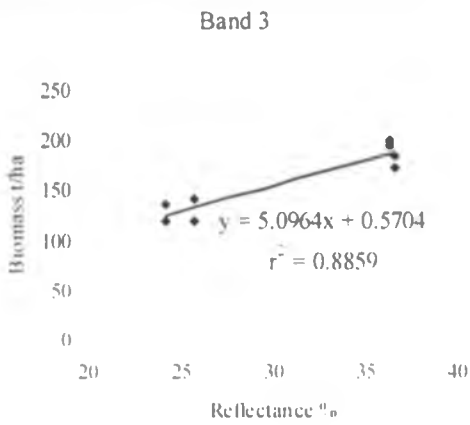
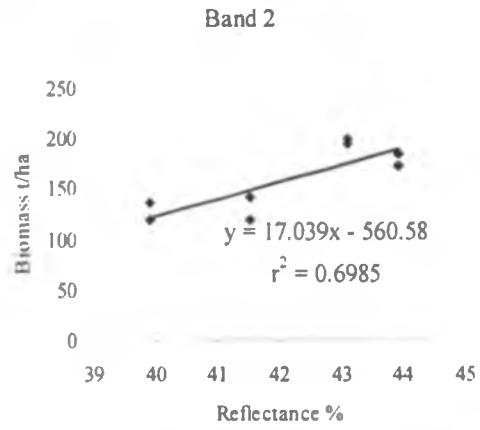
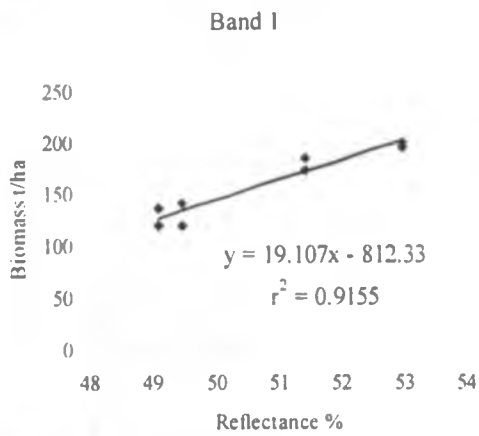


Figure 4-6: ETM+ reflectance vs measured biomass (pixel by pixel) on Field 35

Among the transformed datasets, R/NIR was the best vegetative index for quantifying biomass at senescence stage while NDVI was a better predictor at the crop growing stage. For biophysical parameters, LAI was the best predictor of biomass at both crop senescence and growing stages. However all biophysical parameters and band 4 were negatively correlated with measured biomass at senescence and positively correlated at crop growing stage. At senescence stage (Field 35, Figure 4-4 to Figure 4-9), most of the crop biophysical properties (LAI, f-cover, greenness etc) diminish while biomass increases until it reaches a maximum value. This trend can be explained by the fact that increased senescence leads to reduction of chlorophyll. Chlorophyll reduction is signified by reduced reflectance of band 4. But as senescence progresses, the cane mature more leading to increased accumulation of sucrose and biomass in the stem. This is reflected by the high value of biomass for reduced reflectance value of band 4 signified by the negative correlation shown in Figure 4-4 to Figure 4-6.

At the crop growing stage, Field 34, Figure 4-11 and Figure 4-12), the increase in the biophysical properties of the crop signifies continued crop development with an attendant biomass accumulation. This accounts for the positive correlation at the growing stage.

Generally speaking, the strength of the relationship between the spectral variables and biomass increased as the sugarcane was stratified into more specific pixel by pixel analysis. For example, the relationships between biomass and the spectral variables were generally strongest when the sugarcane field data were evaluated separately for each field pixel by pixel ( $r^2$  values ranged from 0.01 to 0.94) as compared to field by field ( $r^2$  values ranged from 0.001 to 0.82), and when all the sites were combined (36 sites combined,  $r^2$  ranged from 0.05 to 0.2, Table 4-4).

There was a strong relationship between biomass and most spectral variables. The ETM+ visible bands were consistently the most strongly correlated with biomass, with band 1 (blue-green), band 3, and band 4 (near infra-red) being the stronger predictors ( $r^2 = 0.93, 0.85, \text{ and } 0.91$ , respectively). Band 6 was also found to be a good predictor



of biomass ( $r^2=0.86$ ). This corroborates the works of Todd *et al.* (1998) who found band 3 to be the best predictor of biomass for the steppes of eastern Colorado. The predictions were improved further when the analysis was done pixel-by-pixel giving the best prediction of  $r^2$  of 0.98. In comparison, when the sample plots were averaged to cover half a pixel, the prediction was slightly lower ( $r^2 =0.92$ ). The lowest prediction was given when the plots were considered plot by plot ( $r^2 = 0.81$ ).

Total vegetation cover and LAI were also strongly correlated with the spectral data. The reason for the poorer correlation for sample-plot to sample plot parameters compared to the mean normalized values (pixel by pixel) could be attributed to poor geo-referencing of the in-situ measurements at 0.15 pixels or 4.5 m resulting in extracted values representing datasets from fields 4.5 m off the sampling site. This effect is reduced when the entire pixel is considered as a unit with the mean values from the image and the sampling plots quantifying the biomass present more efficiently with a coefficient of efficiency of 0.98 compared to 0.81 for plot by plot. From Table 4-4, the visible bands were most frequently selected as the best predictor variables for Field 34 with active photosynthesis and therefore with higher chlorophyll as compared to Field 35; in which, both the visible and invisible bands were good predictors.

Results made by using various vegetative indices are summarized in Table 4-4. Very high  $r^2$  values of 0.98, 0.98, and 0.99 were obtained for NDVI, SAVI and R/NIR, respectively. However, these  $r^2$  values were less than the values obtained for the models developed with reflectance values in different bands. This implies that at this stage,  $r^2$  values do not justify the use of NDVI over the use of reflectance values of different bands for the development of models for the prediction of plant biomass.

A comparison between the observations taken at different image acquisition dates indicates better results for data gleaned closer to satellite overpass date than that gleaned nearly a month later. It confirms the importance of temporal resolution in

remote sensing as it limits its application to the period of data acquisition (Millinton *et al.*, 1989, Lillesand and Kieffer, 2000).

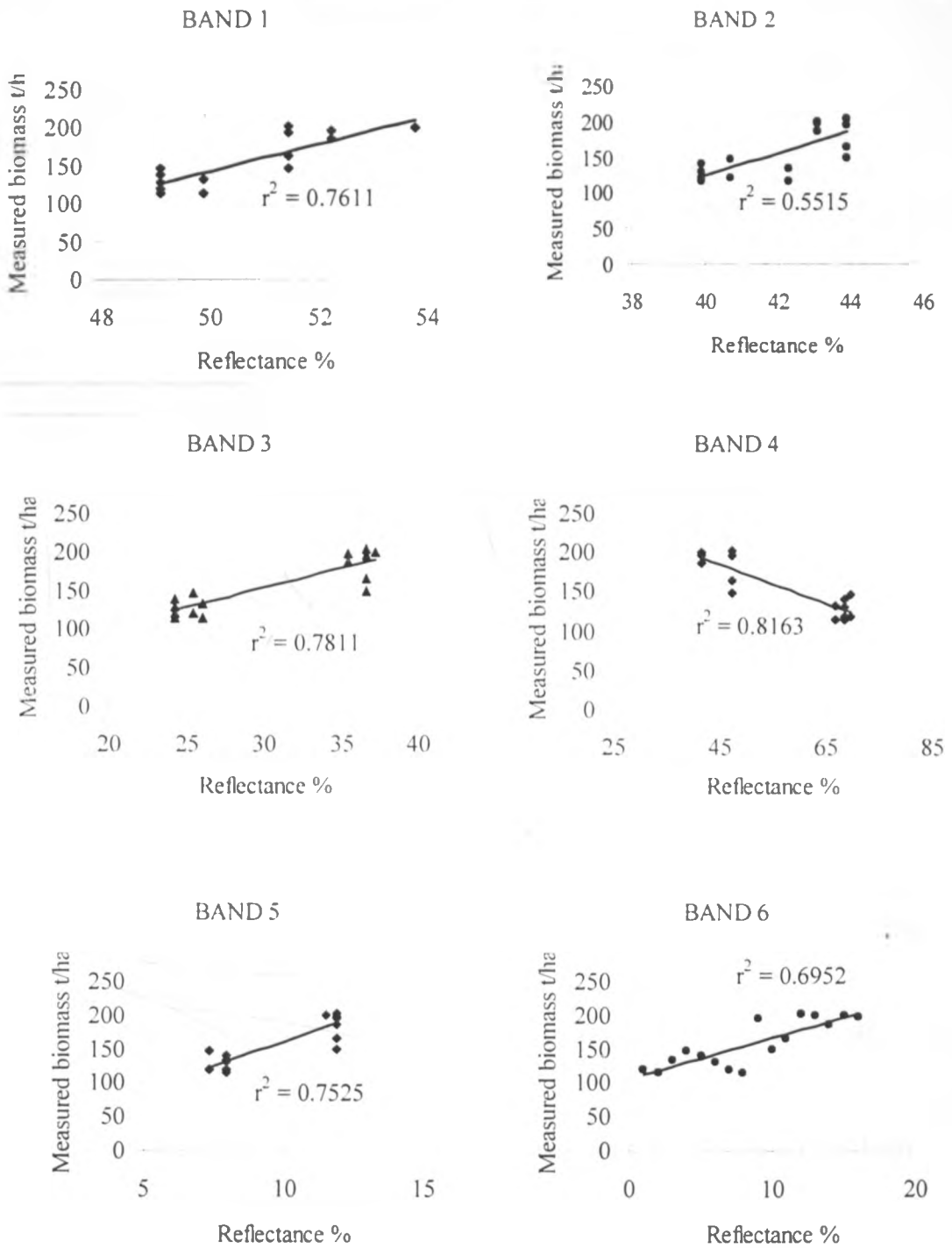


Figure 4-7: ETM+ reflectance vs. measured biomass on all sampling plots (Field 35)

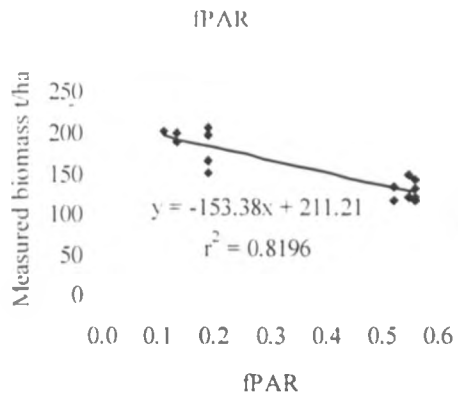
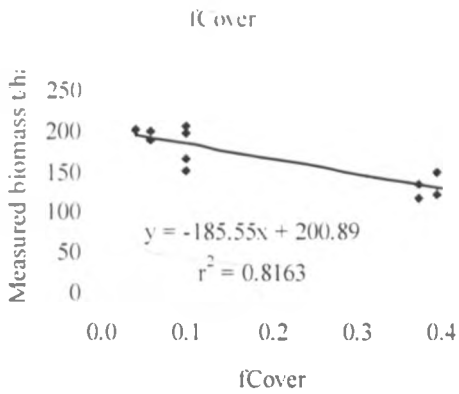
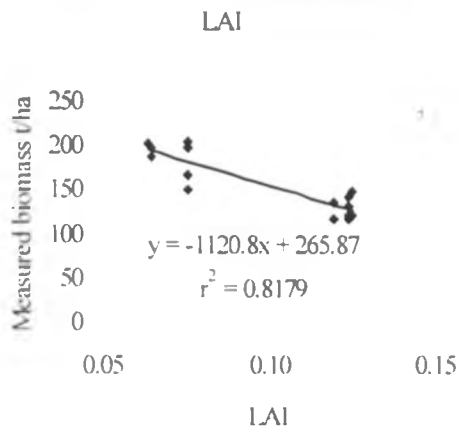
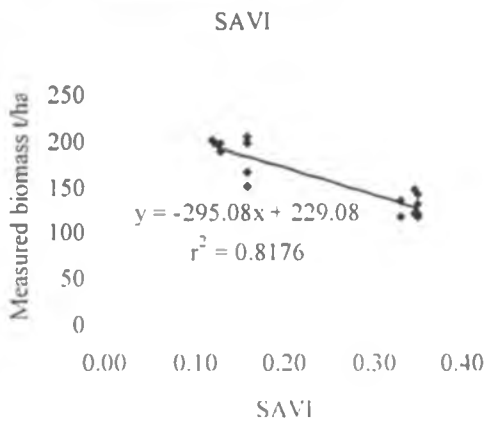
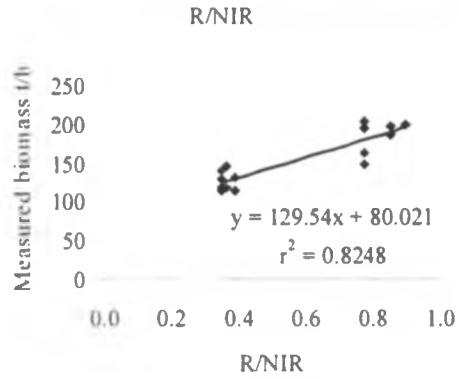
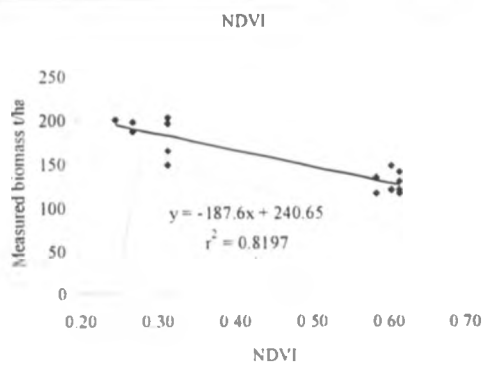


Figure 4-8: Biophysical parameters vs biomass: all sampling plots (Field 35)

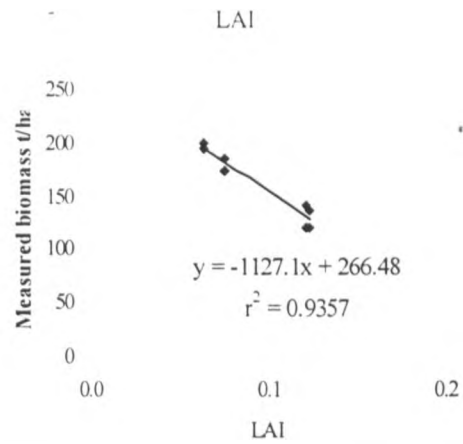
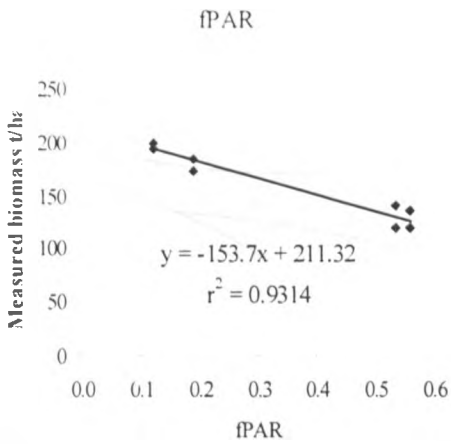
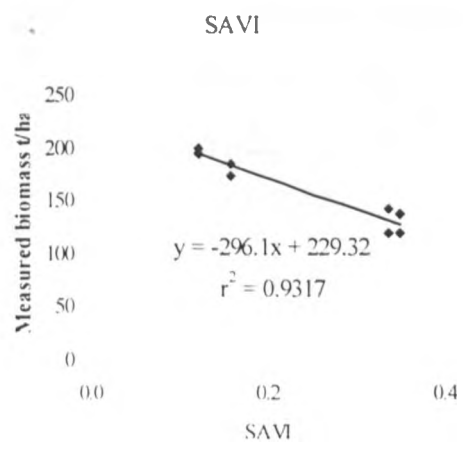
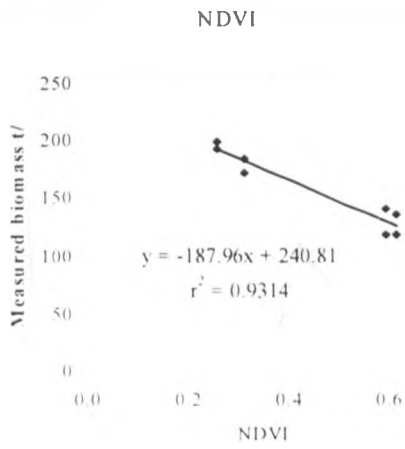
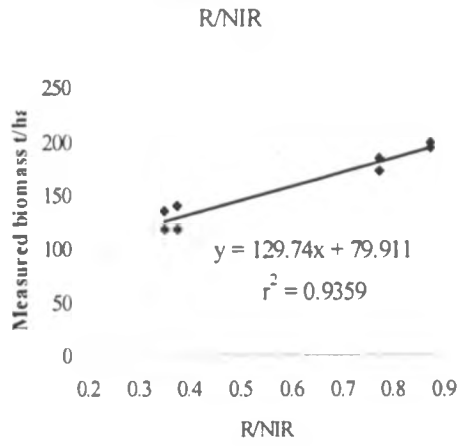
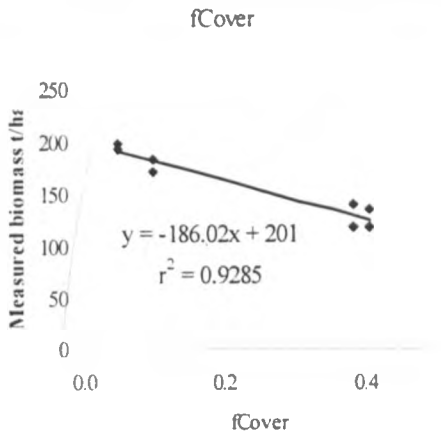


Figure 4-9: Biophysical vs measured biomass pixel by pixel (Field 35)

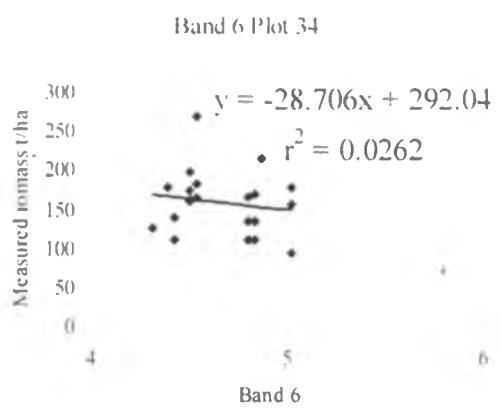
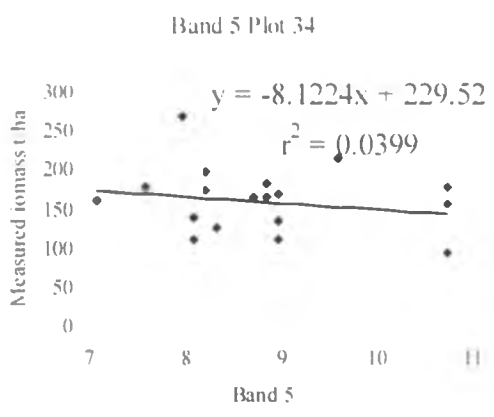
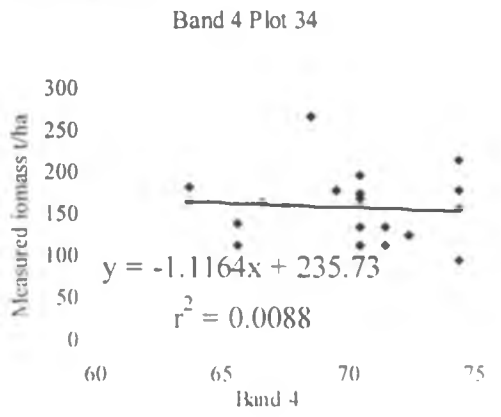
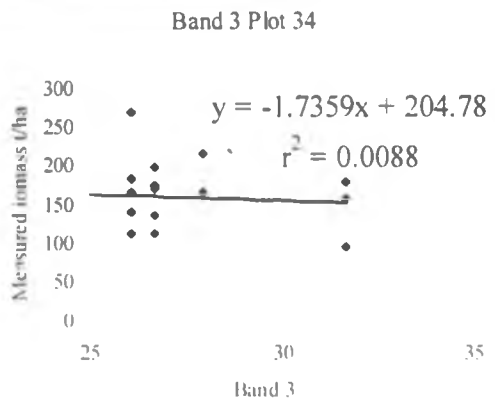
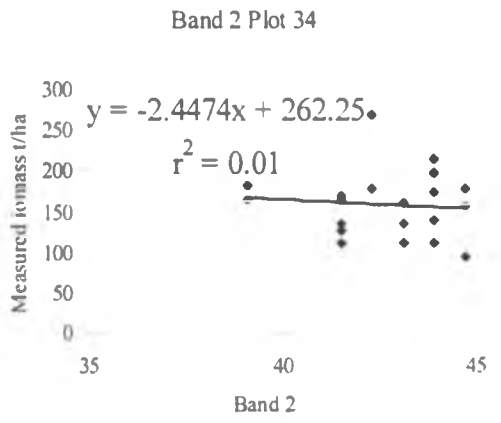
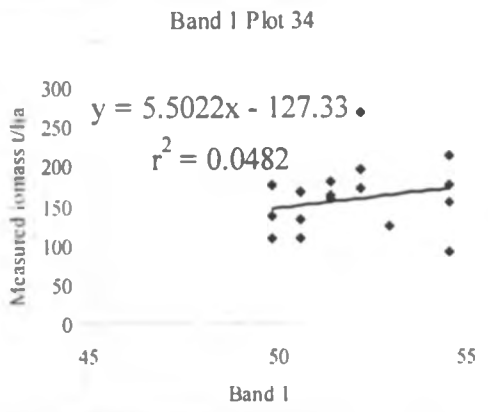


Figure 4-10: Band reflectance vs. measured biomass: all sampling plots (Field 34)

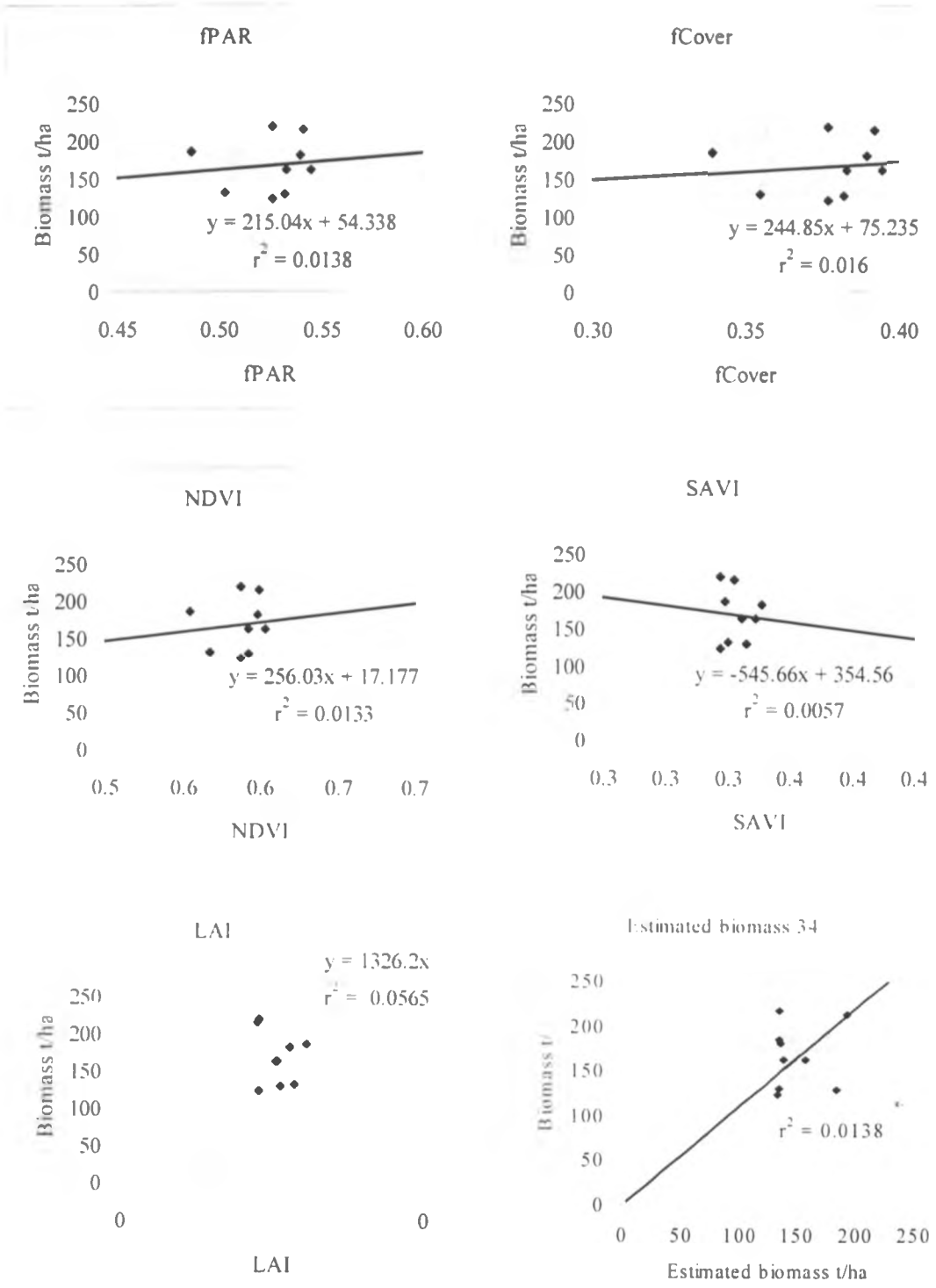


Figure 4-11: Biophysical vs measured biomass bi-pixel by bi-pixel (Field 34)

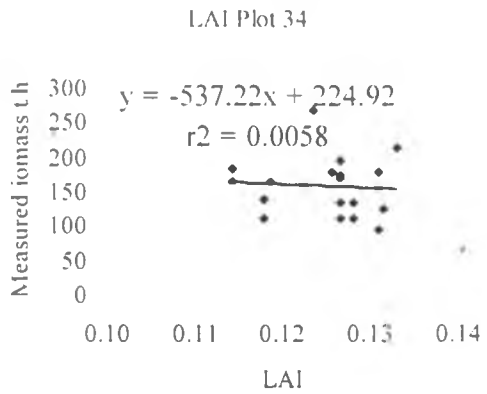
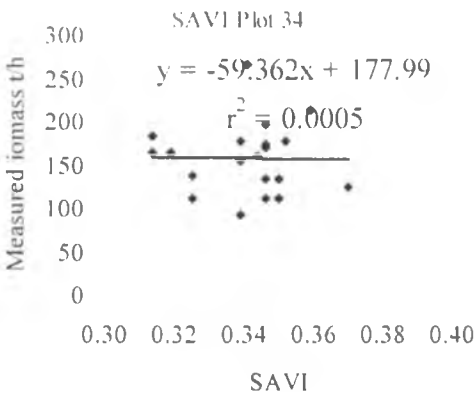
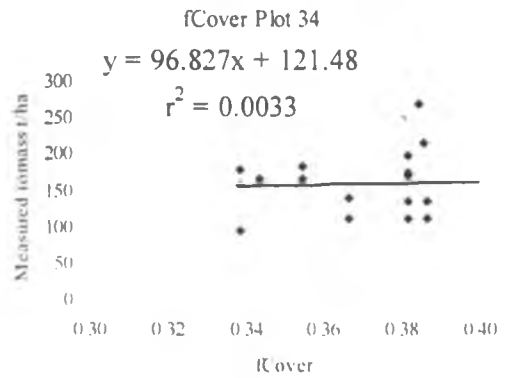
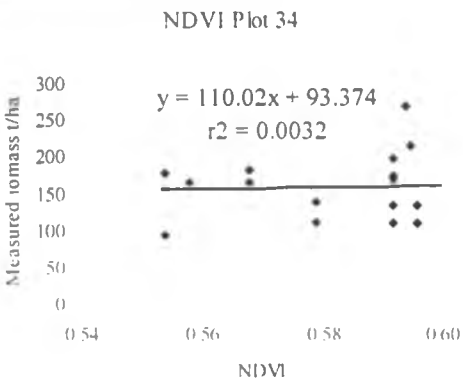
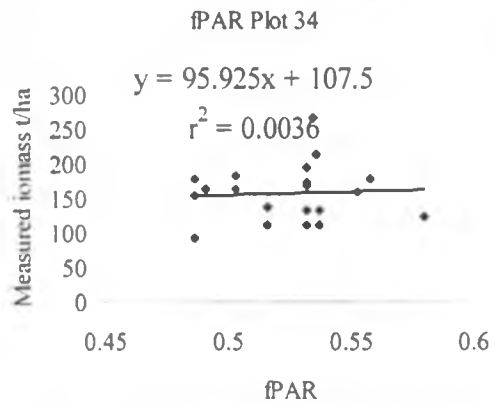
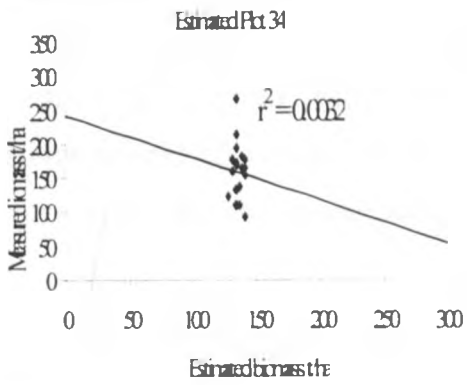


Figure 4-12: Biophysical parameters vs biomass: all sampling plots (Field 34)

A high consistency in the most useful NIR band for inclusion in NDVI for quantifying sugarcane crop biomass was observed. The MIR bands were also found useful for prediction at the senescence stage of the crop growth.

### 4.3.1 Biomass quantification model and mapping

From regression analysis, the following biomass quantification model was developed for estimating biomass using the best biophysical parameter computed by transforming spectral values. The model developed for quantifying biomass yield in t/ha for a sugar cane crop at senescence stage was derived as;

$$Accbiomass = Estbiomass * k - c \quad 4-1$$

Where;

Accbiomass is the predicted biomass in t/ha and

k is the *gain*

c is the *offset*

While Estbioms is the estimated biomass calculated as:

$$Estbioms = \sum_{t=1}^n \epsilon(t) * AcAPAR_{t,}$$

k and c are constants derived as 1.017 and 13.7 for Mumias Nucleus Estate

A script written using the developed model was then run in ILWIS to predict and map biomass using the February 2003 LANDSAT ETM+ image clipped to cover the Mumias Nucleus Estate. The map given in Figure 4-13 is the biomass quantification map for the Estate and surrounding areas showing three levels of production: high (light brown), medium (blue) and sparse or areas without sugarcane crop (greenish). The areas with cane at mature stage and high biomass (more than 160 t/ha) show the light brown color while areas with high chlorophyll at the cane growing stages (120-160 t/ha of cane) are shown with shades of blue. The areas outside the estate were mixed fields showing shades of green while within the Estate, the shades of green are areas with very young cane and bare grounds.

The script determines the crop biophysical properties like LAI, f-PAR, f-cover, NDVI and SAVI and uses the methodology developed by Liang (2004) to estimate and map biomass based on energy fluxes derived from radiation indices extracted from the image.



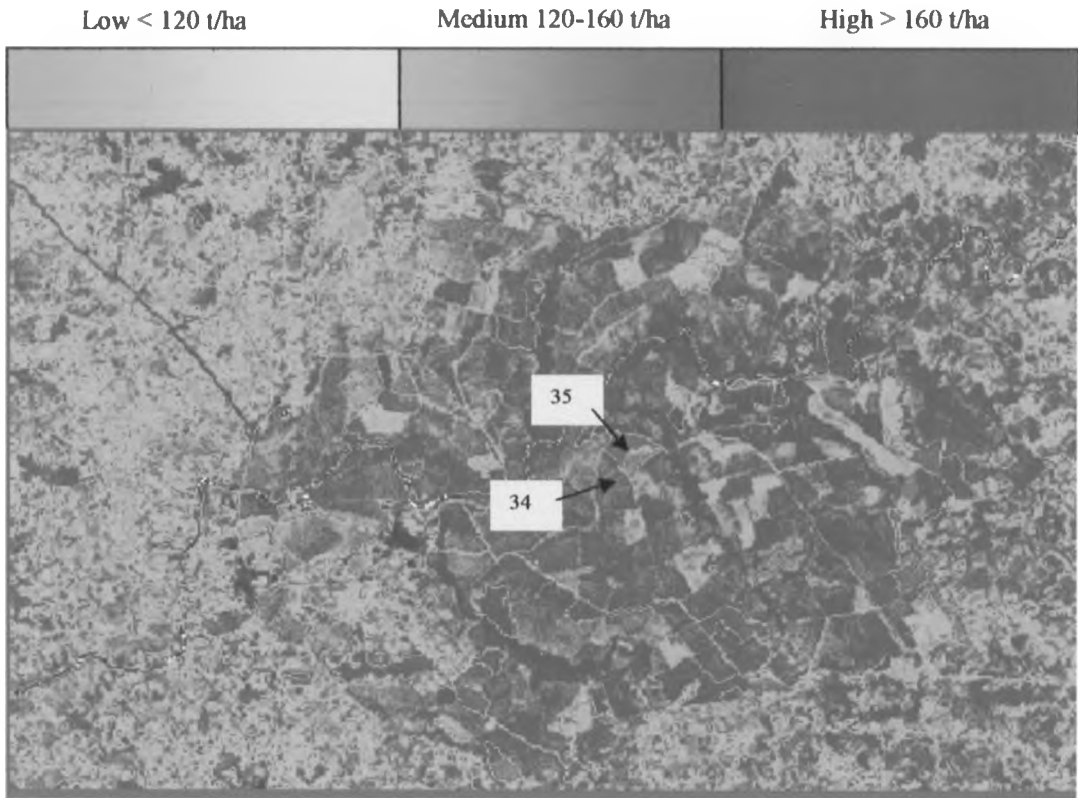


Figure 4-13: Predicted sugarcane biomass for Mumias Nucleus Estate

#### 4.3.2 Model validation

The model validation was conducted by comparing the in-situ measurements and the derived biomass parameter values (Figure 4-14). The comparisons of crop structure parameters showed a high coefficient of correlation ( $r^2$ ) of 0.98 for analysis done pixel by pixel (Blocked). In this case the four datasets gleaned per plot were averaged to normalize them and this was compared with the mean of the extracted spectral or transformed values from the same block. This was repeated for half the block in which case two samples were averaged from each block. The half block comparisons gave a lower correlation value ( $r^2$  of 0.92). Finally, comparisons were made plot by plot that gave the lowest correlation value ( $r^2$  of 0.81). The student t test returned a higher  $r^2$  value of 0.92 for block analysis compared to  $r^2$  of 0.82 for plot-to-plot analysis. This was also reflected in the F test that gave  $r^2$  of 0.91 for block analysis compared to  $r^2$  of 0.53 for plot-to-plot analysis. Randomized blocks reduced the noise.

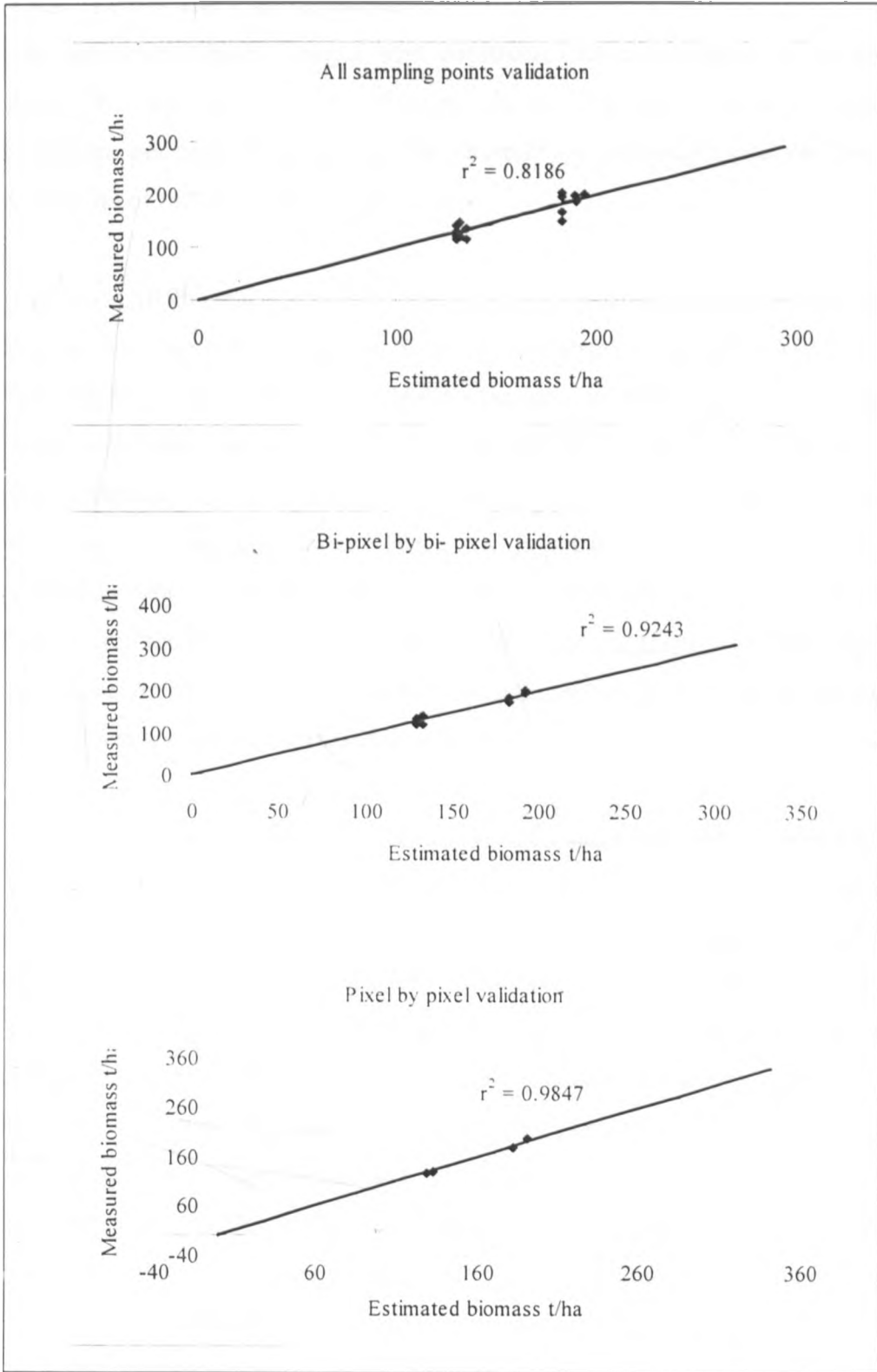


Figure 4-14: Comparison between measured and estimated biomass (Field 35)

### 4.3.3 Errors and Contingencies

The randomized block designs were constructed to reduce noise or variance in the data. This was achieved by dividing the sample into relatively homogeneous subgroups per pixel to ensure that the variability within each block was less than the variability of the population (Table 4-5).

The SSE, ARPE, and  $C_{\text{eff}}$  values indicating the performance of different models are presented in Table 4-5. All predictions show an absolute value for ARPE of no more than 0.02%. From table 4-5, the ARPE for plot -by-plot analysis showed that the model over-predicted the values while when the values were considered pixel by pixel then it showed that the model under-predicted the values. However the value was the same at 0.017 for both cases. These ARPE values (-0.0175 and 0.0175) for the validation data sets indicate a highly satisfactory performance of the model. The SSE was 19 for the pixel-by-pixel analysis with a  $C_{\text{eff}}$  value of 0.98 showing excellent predictive ability. The Jack-Knife cross validation results confirmed that the method used for estimation was unbiased (Table 4-6)

In general, better or comparable predictive ability of models were obtained for most of the parameters, compared to previously reported values in the literature. For example, in a recent study, Thenkabail *et al.* (2000), in their experiment based on a spectral data set acquired from a hand-held spectroradiometer, obtained significant  $r^2$  values for four-band models of 0.78, for LAI. In this study, a value of 0.81-0.99 was found. For the NDVI-based models, the 0.86 got is less than in this study case that gave the highest  $r^2$  value of 0.93. Shanahan *et al.* (2001) reported significant  $r^2$  of about 0.9 between corn yield and broadband NDVI. They acquired images from an airborne platform in four bands over an experiment with varying nitrogen levels. Similarly, in an experiment over different corn hybrids grown under various nitrogen levels, Blackmer and Schepers (1996), using a portable spectroradiometer, reported  $r^2$  values for yield prediction models ranging between 0.70 and 0.99.

Generally, it is difficult to make a true comparison of results with past studies because none of these studies reported estimation from LANDSAT ETM+ data sets acquired over a sugarcane field at senescence stage. Since, the spectral response of vegetation primarily depends on the total vegetation amount, the good relationships for the crop parameters, which are directly dependent on the total foliage cover (LAI, f-Cover, biomass etc.), were generally obtained. It was noted that in the regression models, the spectral values associated with the pixels in the specific parts of an image do not have a one-to-one relationship with the ground-truth measurements. Observations were taken only at a few locations or on a few representative plants in each plot. Complexity is added because each pixel in the image represents a complex response of a ground parcel wherein the proportional coverage of a crop and soil vary within a pixel and also from pixel to pixel. Thus, the variability among pixels pertaining to a particular plot could not be fully associated with the variability in the samples of a given parameter in the plot. This latter problem could lead to a distortion of the relationships and explains the lower correlation values derived for the field as a whole as compared to the higher values correlated for mean normalized plots.

Table 4-5: Performance of prediction and validation models

Parameter		Blocked n=4	Bi-Pixel n=8	All Plot s n=16
SSE		19	79	212
ARPE		-0.017	-0.017	0.017
C <sub>eff</sub>		0.98	0.94	0.81
Mean	Measured	157	140	157
	Estimated	160	142	160
Variance	Measured	1226	1118	1183
	Estimated	1057	906	847
Standard deviation	Measured	35	33	34
	Estimated	33	30	29
Confidence	Measured	25	17	17
	Estimated	23	15	14
Student t test		0.92	0.87	0.82
F test		0.91	0.79	0.53
r <sup>2</sup>		0.98	0.92	0.81

Table 4-6: Jack-knife cross validation results

<b>(Jack-Knife samples) Confirmation of unbiased estimation using Jack-Knife technique on measured biomass</b>																<b>Mean</b>
120	116	133	147	140	130	120	116	196	148	164	203	200	187	200	197	157
	116	133	147	140	130	120	116	196	148	164	203	200	187	200	197	160
120		133	147	140	130	120	116	196	148	164	203	200	187	200	197	160
120	116		147	140	130	120	116	196	148	164	203	200	187	200	197	159
120	116	133		140	130	120	116	196	148	164	203	200	187	200	197	158
120	116	133	147		130	120	116	196	148	164	203	200	187	200	197	158
120	116	133	147	140		120	116	196	148	164	203	200	187	200	197	159
120	116	133	147	140	130		116	196	148	164	203	200	187	200	197	160
120	116	133	147	140	130	120		196	148	164	203	200	187	200	197	160
120	116	133	147	140	130	120	116		148	164	203	200	187	200	197	155
120	116	133	147	140	130	120	116	196		164	203	200	187	200	197	158
120	116	133	147	140	130	120	116	196	148		203	200	187	200	197	157
120	116	133	147	140	130	120	116	196	148	164		200	187	200	197	154
120	116	133	147	140	130	120	116	196	148	164	203		187	200	197	154
120	116	133	147	140	130	120	116	196	148	164	203	200		200	197	155
120	116	133	147	140	130	120	116	196	148	164	203	200	187		197	154
120	116	133	147	140	130	120	116	196	148	164	203	200	187	200		155
																157
<b>Total jack = 16*157-15*mean (t) =157. Hence it is an unbiased estimator</b>																

## 5. Conclusion and Recommendations

### 5.1 Conclusion

The study shows that mature sugarcane biophysical parameters of leaf area index (LAI), fractional cover (f-cover) and photosynthetically active radiation (PAR) could be well predicted in spectral reflectance. It was demonstrated that LAI is the best predictor of sugarcane biomass.

Biomass quantification results from the reported study show that sugarcane biomass can be quantified from single-date LANDSAT ETM+ crop reflectance for large sugarcane fields.

The results remained positive for comparisons across the field for predicted values and measured values of mature sugarcane field. In general, a good agreement between the observed and predicted values of various parameters was observed. When the VI-based models for estimating biomass measurements were compared to the single reflectance band-based models, the former were found to be better, thus making a case for the use of VI-based models.

The results advocate the appropriateness of satellite imagery observations for development of models for estimation of sugarcane crop parameters and yield. This model could further be used for either the direct estimation of sugarcane crop condition indicating parameters, or they could be linked to various physical crop growth models. The application of the results at the field scale could be visualized for the development of crop conditions and/or application maps.

A potential application for this research is in conjunction with cane inventory assessments at regular intervals throughout the milling season to provide area-based estimates of the standing crop, as well as for quantifying how long the crop is likely to remain in the field before being harvested. This information is particularly important in the Kenyan sugar zone context, where climatic variability has large impacts on cane

supply and where energy from grid supplies is expensive. This information could assist mill and cane supply managers to optimize sugar production and biomass supply (Figure 5-1) for co-generation through improved cane supply scheduling, given a yield estimate and the age distribution of the standing crop

## 5.2 Recommendations

1. To make the results more applicable, further research is needed to develop relationships under the influence of other factors, such as soil moisture, organic matter, and crop residue.
2. Further research could also be done to establish models for predicting sugarcane phenological development and crop variety classification.
3. There is need also to establish the efficacy of using satellite imagery on sugarcane fields by testing images from different satellites like SPOT, Quick Bird and EROS



Figure 5-1: Sugarcane fields (harvested and growing)

## References

- Albert, J.P. Elizabeth, A. Walter-Shea. Leiji. Andres,V. Hayes, M and Groboda, M.D. (2002).** Drought monitoring with NDVI-based standardized vegetation index. *Journal of Photogrammetry and Remote Sensing, American Society of Photogrammetry and Remote Sensing*, volume 68, No1, pp71-75.
- Asrar, G. Mynenei, R.B and Choudhury, B.J. (1992).** Spatial heterogeneity in vegetation canopies and remote sensing of absorbed photosynthetically active radiation: A modelling study, *Remote Sensing of Environment* 41:86-103.
- Asrar, G. Kanemasu, E.T and Yoshida, M. (1985).** Estimation of leaf area index from spectral reflectance of wheat under different cultural practices and solar angle. *Remote Sensing of Environment*. 17(1): 1-11.
- Batchly, K and VanLeeuwen, W. (1997).** Comparison of vegetation indices global set of TM Images for EOS-MODIS. *Remote Sensing of Environment*, 59, 440-451.
- Bausch, W. C. (1993).** Soil background effects on reflectance-based crop coefficients for corn. *Remote Sensing of Environment* 46:213-222.
- Betrann, M.C and Belmonte, A.C. (2001).** Irrigated crop area estimation using LANDSAT imagery in La Mancha Spain, *Journal of Photogrammetric Engineering and Remote Sensing, American society of Photogrammetry and Remote Sensing*, Volume 67(10), pp1177-1184.
- Blackmer, T.M and Schepers, J.S. (1996).** Aerial photography to detect nitrogen stress in corn. *Journal of Plant Physiology* 148:440–444.
- Brown, R.J. Staenz, H. McNairn, H. Hopp, B and van Acker, R. (1997).** Application of high-resolution optical imagery to precision agriculture. Presented at International Symposium, Geomatics in the Era of RADARSAT (GER '97). Ottawa, Canada: Canadian Institute of Geomatics, the Canadian Aeronautics and Space Institute, the Canadian Space Agency, and the Canada Centre for Remote Sensing, Natural Resources Canada.



- Brown, R. Slater, J and Askew D. (1987).** Integrating satellite imagery and GIS: The Experience of a BNSC funded project to monitor environmentally sensitive areas, Technical Applications, pp 15-21.
- Chavez, P.S. (1996).** Image based atmospheric correction-revised and improved, Photogrammetric Engineering and Remote Sensing, 62(9), 1025-1036.
- Choudhury, B. J. (1994).** Synergism of multispectral satellite observations for estimating regional land surface evaporation. Remote Sensing of Environment 49:264-274.
- Cloutis, E.A. (1999).** Agricultural crop monitoring using airborne multi-spectral imagery and C-TM synthetic aperture radar, International Journal of Remote Sensing, Vol. 20, no. 4, 767± 787.
- Congalton, R.G. (1986).** Accuracy assessment and validation of remotely sensed and other spatial information. International Journal of Wildland Fire, Volume 10 (Issue 4) Pages 321-328.
- Congalton, R.G. Balogh, M.C. Bell, K and Green, J.A. (1998).** Mapping and monitoring agricultural crops and other land cover in the Lower Colorado River Basin, Photogrammetric Engineering and Remote Sensing.
- Crist, E. P. (1984).** Comparison of coincident Landsat-4 MSS and TM data over an agricultural region. Technical Papers of the 50th Annual Meeting of the American Society of Photogrammetry, Washington, D.C., pp. 508-517.
- Crist, E.P. Laurin, R and Cicone, R.C. (1986).** Vegetation and soils information contained in transformed Thematic Mapper data, Proceedings of IGARSS' 86 Symposium, 1465-70. Ref. ESA SP-254. Paris: European Space Agency.
- Curran, P.J. Kupiec, J.A and Smith, G.M. (1997).** Remote sensing the biochemical composition of slash pine canopy. IEEE Trans. Geoscience Remote Sensing 35(2): 415-420.
- D'Souza, G. Wooster, M. Bonifacio, R and Drake, N. (1995).** Basic Remote Sensing Principles and Background (Surlock JMO, Wooster MJ and D'Souza G, Ed) Version 1 Natural Vegetation as a Resource: A Remote Sensing workbook for East and Southern Africa, Commonwealth Science council, London. pp A.10-15.

- Fraysse, G. (1980).** Remote sensing application in agriculture and hydrology Institute of Agrophysics Polish Academy of Sciences.
- Frouin, R and Pinker, R. T. (1995).** Estimating photosynthetically active radiation (PAR) at the earth's surface from satellite observations. *Remote Sensing of Environment* 5 1:98-107
- GARMIN. (1998).** GPS 12XL, Personal Navigator, Owners manual and Reference, Garmin (Europe) Ltd.
- Government of Kenya. (1971).** Topographical map for Mumias area, 102.2
- Guo, X. Price, K.P and Stiles, J.M. (2000).** Modeling biophysical factors for grasslands using LANDSAT TM data in Eastern Kansas, annual meeting, The Second International Conference on Geospatial Information in Agriculture and Forestry. Lake Buena Vista, Florida. January 10-12.
- Hall, F. G. Huenuenrich, K. F. and Goward, S. N. (1990).** Use of narrow-band spectra to estimate the fraction of absorbed photosynthetically active radiation. *Remote Sensing of Environment* 28:47-54.
- Hankins, M. (1989).** Renewable energy in Kenya. Mortif creative arts, Nairobi.
- Hardy, E.E. (1989).** Remote sensing for land and water resource management, Outgrowth of National congress on remote sensing for resource management. Soil Conservation Society of America pp 3.
- Hartl, Ph. (1989).** Fundamentals of remote sensing Applications of remote sensing in agro-meteorology. Proceedings of a course held by the joint Research Centre of the commission of European communities in the framework of the ISPRA-courses. ISPRA , Italy. Pp 1-18.
- Hay, S. Rogers, J.D and Wint, W.G.R. (1997).** Natural Resource Assessment for Nigeria: Using Meteorological satellite data to map land cover, Technical Applications, pp 23.
- Holben, B.N. Tucker, C.J and Fan, C.J. (1980).** Spectral assessment of soybean leaf area and leaf biomass, *Photogrammetric Engineering and Remote Sensing*, 46, 651-656.
- Huete, A.R. (1988).** A Soil Adjusted Vegetation Index (SAVI) *Remote Sensing of the Environment* Volume 25, pp. 295-309.

**ILWIS 3.2 academic. (2005).** ITC, The Netherlands.

**Inman-Bamber, N.G. (1994).** “Temperature and seasonal effects on canopy development and light interception of sugarcane”, *Field Crops Research*, 1994, 36 pp 41-51

**Inoue, Y and Morinaga, S. (1995).** Estimating spatial distribution of plant growth in a soybean field based on the remotely sensed spectral imagery measured with a balloon system. *Japanese Journal. Crop Science* 64(1): 156-158.

**Inoue, Y. Morinaga, S and Tomita. A. (2000).** A blimp-based remote sensing system for low-altitude monitoring of plant variables: A preliminary experiment for agriculture and ecological applications. *International Journal Remote Sensing* 21(2): 379-385.

**International Centre for Research in Agro forestry, ICRAF. (2000).** Improved land management in the Lake Victoria basin: linking land and lake, research and extension, catchment and lake basin, Final Technical Report, start-up phase July 1999 to June 2000, pp 1.

**Isaaks, E. H and Srivastava, R.M. (1989).** An introduction to applied geostatistics. Oxford Univ. Press, New York, Oxford.

**Jaetzold, R and Schmidt, H. (1982).** Farm Management Handbook of Kenya. Conditions and Farm Management Information. Volume .II - Part 1 - 3 (A=West Kenya;).

**Jago, R.A. Cutler, M.E.J and Curran, P.J. (1999).** Estimating canopy chlorophyll concentration from field and airborne spectra. *Remote Sensing of Environment*, Volume 68, Number 3, pp. 217-224(8), Elsevier Science

**James, L.D and Burges, S.J. (1982).** Selection, calibration, and testing of hydrologic models, *Hydrologic modeling of small watersheds*. American Society of American Engineers. St. Joseph, Michigan, pp.437-472.

**Jensen, J.R. (1996).** Introductory digital image processing, ARS perspective, 2nd edition. Prentice Hall Inc. NJ USA Chapter 1 pp1-8.

**Kenya Sugar Board (KSB). (2003).** Annual Bulletin.

**Kenya Sugar Board (KSB). (2002).** Annual Bulletin

- Kituyi, E. Marufu, L. Huber, B. Wandiga, S.O. Jumba, S.O. Andreae, M.O. Helas, G. (2001a).** Biofuel consumption rates and patterns in Kenya, *Biomass and Bio energy Journal*, Pergamon.
- Kituyi, E. Marufu, L. Wandiga, S.O. Jumba, S.O. Andreae, M.O and Helas, G. (2001b).** Biofuel availability and domestic use patterns in Kenya, *Biomass and Bio energy Journal*, Pergamon.
- Klemin, J and Fagerlund E. (1987).** Influence of different nitrogen and irrigation treatments on the spectral reflectance of barley. *Remote Sensing of Environment*
- Liang, S. (2004).** Quantitative remote sensing of land surfaces (New Jersey), John Wiley and Sons.
- Lillesand, T.M and Kiefer, R.W. (2000).** Remote sensing and image interpretation 4<sup>th</sup> Edition, John Wiley and Sons inc. USA.
- Maas, S.J. (1998).** Estimating Cotton Canopy Ground Cover from Remotely Sensed Scene Reflectance, *Agronomy Journal*, - agron.scijournals.org Page 1.
- Malingreau, J.P. (1989).** Vegetation index and the study of vegetation dynamics, Proceedings of a course held by the joint research centre of the commission of European Communities in the framework of the ISPRA-courses, Italy pp 285-301.
- Matsuda, S. Kubota, H and Iwaki, H. (1982).** Biomass energy: Its possibility and limitation. *Science* 52. 735-741.
- Millington, A and Townseed. (1989).** Biomass Assessment, Woody biomass in the SADCC region.
- Ministry of Energy. (1992).** Government of Kenya, Survey on Recurrent and Renewable Energy Systems for Rural Development in Kenya.
- Munden, R. Curran, P.J and Catt, J.A. (1994).** The relationship between red edge and chlorophyll concentration in the broadbalk winter wheat experiment at Rothamsted, *International Journal of Remote Sensing* 15 (1994), pp. 705–709.

- Neale, C. M. U. Bausch, W. C and Heeremann, D. F. (1989).** Development of reflectance-based crop coefficients for corn. *Transactions of the ASAE* 32(6): 1891-1899.
- Nyang, F.O. (1998).** Household Energy Demand and Environmental Management in Kenya. Publised PhD Thesis, Universiteit van Amsterdam, Netherlands.
- Omuto, C.T. (2003).** Rapid mapping of hydraulic conductivity in a tropical watershed. MSc.thesis (Department of Agricultural Engineering). University of Nairobi. [www.uonbi.ac.ke/dept](http://www.uonbi.ac.ke/dept)
- Parodi, Ir. G. (2000).** AVHRR Hydrological Analysis System – Algorithms and theory, Version 1.0. WRES, ITC, the Netherlands.
- Patel, N.K. Patnail, C. Dutta, S. Shekh, A.M and Dave, A.J. (2001).** Study of crop growth parameters using airborne imaging spectroradiometer. *International Journal of Remote Sensing* 22(12): 2401-2411.
- Priestley, C.H.B and Taylor, R.J. (1972).** "On the assessment of surface heat flux and evaporation using large scale parameters." *Mon. Wea. Rev.*, **100**, 81-92.
- Prince, S.D. (1991).** Satellite remote sensing of primary production: comparison of results for Sahelian grasslands 1981± 1988. *International Journal of Remote Sensing*. 12: 1301± 1311.
- RCSSMRS. (1988).** Regional Centre for Services in Surveying, mapping and remote sensing. Earth Resources Mapping in Africa, Nairobi.
- Shanahan, J.F. Schepers, J.S. Francis, D.D. Varvel, G.E. Wilhelm, W.W. Tringe, J.M. Schlemmer, M.R and Major, D.J. (2001).** Use of remote-sensing imagery to estimate corn grain yield. *Agronomy Journal*. 93(3): 583-589.
- Skidmore, A.K. Bijker, W. Schmidt, K and Kumar, L. (1997).** Use of Remote Sensing and GIS for Sustainable land management, *ITC Journal*, pp 302
- Thenkabail, P.S. Smith, R.B and Pauw, E.D. (2000).** Hyperspectral vegetation indices and their relationships with agricultural crop characteristics, *Remote Sensing of Environment*. 71(2): 158-182.
- Trochim, W.M.K. (2002).** Research methods knowledge base.

- Van der Kwast, J and Zomer, R. (2002).** Land cover classification and DEM generation using SPOT and LANDSAT images in the lake Victoria Basin, Western Kenya, ICRAF and Universiteit Utrecht, Netherlands.
- Van Dillewijn, C. (1952).** Botany of Sugarcane, Waltham, MA, USA.
- Villa Nova, N.A. Pedro Jr, M.J. Pereira, A.R and Ometto, J.C. (1972).** Estimativa de Graus-dia, acumulados acima de qualquer temperatura base, em função das temperaturas máxima e mínima. Caderno de Ciências da Terra, Instituto Geográfico-USP, n.30,1972.
- Warren, S.D. Johnson, M.O. Goran, W.D and Diering V.E. (1990).** An automated, objective procedure for selecting representative field sample sites, article in Journal of Photogrammetric Engineering and remote sensing, American society for Photogrammetry and Remote Sensing, Volume 56(3), pp333-335.
- Wiegand, C.L. Richardson, A.J and Nixon, R.J.N. (1992).** Spectral component analysis: A bridge between spectral observations and agrometeorological crop models. IEEE Transactions, Geoscience and Remote -24: 83-88.
- Woldu, T. (1997).** Remote sensing of biomass production, radiation distribution, and rainfall patterns, senior thesis report, Physics department, University of Asmara.
- Woods, J and Hall, D.O. (1994).** Bioenergy for development, Rome: FAO.  
[www.geografforlaget.dk/course/ENGLISH/basic/tele1.htm](http://www.geografforlaget.dk/course/ENGLISH/basic/tele1.htm)

### **Further Bibliography**

- Arnon, D. I. (1949).** Copper enzymes in isolated chloroplasts polyphenoloxidase in *Beta vulgaris*, Plant Physiology 24(1): 1-15.
- Bianchi, G. (1987).** Role and Strategies of the ENEA in the Biomass Energy Sector.
- Deutschi, C.V and Journel, A.G. (1992).** GSLIB. Geostatistical software library and user's guide. Oxford Univ. Press, Oxford.
- Feroro, G.L. Guassi, G and Williams, H.E. (1987).** Biomass Energy: From Harvesting to Storage (selected papers), Commission of the European Communities, Brussels and Luxembourg.

- Hall, F.G. Shimabukuro, Y.E and Huemrich, K.F. (1995).** Ecological applications, Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. Forest.
- Hauser, L.G. (1980).** Energy: current supplies and future technology; Agricultural energy 1980 ASAE National energy symposium, Voll, Solar energy, ASAE publication 3-81 pp 15-19.
- Heppel, C.M. Worrall, F. Burt, T.P and William, R.J. (2002).** A classification of drainage and macropore flow in agricultural catchment, Hydrological Processes Journal pp16-46.
- Howes, M. (1985).** International Development Research Centre. Manuscript Report, Rural Energy Surveys in the Third World.
- Huete, A.R. Jackson, R.D and Post, D.F. (1985).** Spectral response of a plant canopy with different soil backgrounds, Remote Sensing the Environment Volume 17, pp. 37-53.
- Huete, A.R. (1989).** Soil influences in remotely sensed vegetation canopy spectra. In 'Theory and Application of Optical Remote Sensing'. (G. Asrar, Ed). Wiley, New York pp. 107-141.
- Janssen, L.L.F and Huurneman, G.C. (2001).** Principles of remote sensing, The International Institute of Aerospace Survey and Earth Sciences (ITC), Enschede, The Netherlands.
- Krishnan, P, Alexander, J.D, Butler, B.J and Hummel, J.W. (1980).** Reflective technique for predicting soil organic matter, Soil Science Society of America Journal 44,1282-5.
- Liebhold, A. M. Rossi, R. E and Kemp W. P. (1993).** Geostatistics and geographic information systems in applied insect ecology. Annual Revised Entomology. 38: 303-327.
- Nelson, R.F. (2000).** Secondary forest age and tropical forest biomass estimation using Thematic Mapper imagery, Bioscience.
- Rossi, R.E. Mulla, D.J. Journel, A.G and Franz, E.H. (1992).** Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs: 62: 277-314.

**Tsuyoshi, A. (1999).** Applications of Remote sensing Technology for agro-environmental issues, Proceedings of the 6th JIRCAS International Symposium. GIS applications for Agro-environmental issues in developing regions.

**Zhenkul, Ma. Hart, M.M and Redmond, R.L. (2001).** Mapping vegetation across large geographic areas, integration of remote sensing and GIS to classify multi resource data, Journal of Photogrammetric Engineering and Remote Sensing, American Society for Photogrammetry and remote sensing, Vol 67(3) pp 275-307



## **Glossary of terms**

**Geographic Information Systems (GIS):** is a system of hardware and software used for storage, retrieval, mapping, and analysis of geographic data. It allows for spatial data and associated attributes in the same coordinate system to be layered together for mapping and analysis.

**Global Positioning System:** Is a precise satellite based navigation location system that has a fleet of more than 24 communication satellites that transmit signals globally which through the GPS receivers aids in accurately determining the latitude, longitude and altitude of a position on or above the earth's surface.

**LANDSAT ETM+:** Imagery has a panchromatic band which is able to produce panchromatic images at 12.5 m resolution allowing for panchromatic sharpened multispectral images to be created without rectifying images to one another.

**LANDSAT MSS:** The Landsat Multispectral Scanner (MSS) was a sensor onboard Landsats 1 through 5 and acquired images of the Earth nearly continuously from July 1972 to October 1992, with an 18-day repeat cycle for Landsats 1 through 3 and a 16-day repeat cycle for Landsats 4 and 5.

**Spectral Vegetation Index: (SVI)** is generated by combining data from multiple spectral bands into a single value. Usually simple algebraic formulations, SVIs are designed to enhance the vegetation signal in remotely sensed data and provide an approximate measure of live, green vegetation amount.

**Photosynthetically Active Radiation:** is Radiation between 0.4  $\mu\text{m}$  and 0.7  $\mu\text{m}$  used by the green canopy in the photosynthetic process.

**Greenness:** The difference between normalized near infrared (0.7-1.1  $\mu\text{m}$ ) and visible (0.5-0.7  $\mu\text{m}$ ) radiances of vegetation representing the state of growth of a crop.

## Appendices

### A.1 Script for estimating biophysical parameters

#### Script for estimating biophysical parameters and quantifying accumulated

#### biomass

```
// CALCULATION OF BIOPHYSICAL PARAMETERS USING BIOPHYSICAL//
```

```
// the input variables are maps, and conversion constants//
```

#### Run biophysical

```
// the output of this script are reflectance maps, NDVI, SAVI, LAI, and surface  
temperature map//
```

```
// these outputs are used in climate script to determine climate factors and calculates  
energy balance//
```

```
//CALCULATE ENERGY FLUXES USING SUB-SCRIPT ENERGY//
```

```
// the inputs are from biophysical sub-script and a and b constants from excel iteration
```

```
Run energy
```

```
// the output is evapotranspiration fraction map that is used in the next script
```

```
// CALCULATION OF DAILY EVAOPTRANSPIRATION RATES USING ET24  
SUB-SCRIPT
```

```
Run ETday
```

```
Open Rn24.mpr -noask
```

```
Open evapfraction.mpr -noask
```

```
Open ET24.mpr -noask
```

#### Script for calculations of sensible heat in the spreadsheet

```
//calculating energy fluxes//
```

```
Lincome=(1.24*(appemaETM+(65.30)/(30+273))^(1/7)*5.67e-  
8*(273+30)^4)*longitude/longitude
```

```
Loutcome=surfemis*5.67e-8*surfTemp^4+(1-surfemis)*5.67e-  
8*(273+30)^4*(appemaETM+(65.30)/(30+273))^(1/7)
```

```
Lnet=Lincome-Loutcome
```

```
NetRad=(1-r0)*Kexo+Lnet
```

```

soilheat=NetRad*(SurfTemp-273.15)/(100*r0)*(0.32*1.1*r0+0.62*(1.1*r0)^2)*(1-
0.978*NDVI^4)
zom=exp(-5.5+5.8*NDVI)
zoh = 0.1*zom
displace=1*(1-(1-exp(-
sqrt(20.6*iff(LAI<0,0.00001,LAI)))/(sqrt(20.6*iff(LAI<0,0.00001,LAI))))
ublend = 3.0*(ln(100-displace)-ln(zom))/(ln(5-displace)-ln(zom))
//Calculate the temperature difference //
deltat=iff(0.46608*surfTemp-129.893<0.001,0.001,0.46608*surfTemp-129.893)
//Iteration steps to determine sensible heat flux)
//assume psi-m = 0 and psi-h = 0//
// ITERATION-1//
ustar1=0.41*ublend/(ln((100-displace)/Zom)-0)
resisttah1=(ln((5-displace)/zoh)-0)/(0.41*ustar1)
senheat1=(1.2*1005*deltat)/resisttah1
obulength1=-abs(ustar1^3*1.2*1005*surfTemp/(0.41*9.81*senheat1))
xm1=(1-(16*(100-displace)/obulength1))^0.25
xh1=(1-(16*(5-displace)/obulength1))^0.25
psim1=2*ln((1+xm1)/2)+ln((1+xm1^2)/2)-2*atan(xm1)+0.5*pi
psih1=2*ln((1+xh1^2)/2)
// ITERATION-2//
ustar2=0.41*ublend/(ln((100-displace)/Zom)-psim1)
resisttah2=(ln((5-displace)/zoh)-psih1)/(0.41*ustar2)
senheat2=(1.2*1005*deltat)/resisttah2
obulength2=-abs(ustar2^3*1.2*1005*surfTemp/(0.41*9.81*senheat2))
xm2=(1-(16*(100-displace)/obulength2))^0.25
xh2=(1-(16*(5-displace)/obulength2))^0.25
psim2=2*ln((1+xm2)/2)+ln((1+xm2^2)/2)-2*atan(xm2)+0.5*pi
psih2=2*ln((1+xh2^2)/2)
//ITERATION-3//
ustar3=0.41*ublend/(ln((100-displace)/Zom)-psim2)

```

```

resisttah3=(ln((5-displace)/zoh)-psih2)/(0.41*ustar3)
senheat3=(1.2*1005*deltat)/resisttah3
obulength3=-abs(ustar3^3*1.2*1005*surfTemp/(0.41*9.81*senheat3))
xm3=(1-(16*(100-displace)/obulength3))^0.25
xh3=(1-(16*(5-displace)/obulength3))^0.25
psim3=2*ln((1+xm3)/2)+ln((1+xm3^2)/2)-2*atan(xm3)+0.5*pi
psih3=2*ln((1+xh3^2)/2)
//ITERATION-4//
ustar4=0.41*ublend/(ln((100-displace)/Zom)-psim3)
resisttah4=(ln((5-displace)/zoh)-psih3)/(0.41*ustar4)
senheat4=(1.2*1005*deltat)/resisttah4
obulength4=-abs(ustar4^3*1.2*1005*surfTemp/(0.41*9.81*senheat4))
xm4=(1-(16*(100-displace)/obulength4))^0.25
xh4=(1-(16*(5-displace)/obulength4))^0.25
psim4=2*ln((1+xm4)/2)+ln((1+xm4^2)/2)-2*atan(xm4)+0.5*pi
psih4=2*ln((1+xh4^2)/2)
senheat = ifundef(senheat4,0,senheat4)
//Calculates the latent heat flux map//
latheat=iff(NetRad-Soilheat-senheat<0.0,NetRad-Soilheat-senheat)
evapfraction { dom=value.dom;vr=0.000:1.00:0.001 }=latheat/(NetRad-soilheat)
Script for calculation of LATitude AND LONGitude MAPS to project the coefficients
in same mapbase
//calculation of LATitude AND LONGitude MAPS to project the coefficients in same
mapbase//
latitude { dom=value.dom;vr=-180:180:0.00001 }:=
iff(B03_1.crdy(transform(mapcrd(B03_1),latlon)),0)
longitude { dom=value.dom;vr=-180:180:0.00001 }:=
iff(B03_1.crdx(transform(mapcrd(B03_1),latlon)),0)
lat { dom=value;vr=0:24:0.0001 }:=((10.45-(4*(45-longitude)/60)+et(DA(41))/60)
omeg=15*(lat-12)*pi/180
coszen { dom=value;vr=-1:1:0.0001 }:=cos((90-59.3)*pi/180)*iongitude/longitude

```

```

Kexo{dom=value;vr=-100:10000:0.0001}:=1367*E0(da(41))*coszen
// RADIOMETRIC CALIBRATION //
//radiance =Lmin + {(Lmax-Lmin)/(Qcalmax-Qcalmin)}*(DN-Qcalmin)
L141{dom=value.dom;vr=-200:10000:0.001}:-6.2+((191.6-(-6.2))*(B03_1-1)/254)
L241{dom=value.dom;vr=-200:10000:0.001}:-6.4+((196.5-(-6.4))*(B03_2-1)/254)
L341{dom=value.dom;vr=-200:10000:0.001}:-5.0+((152.9-(-5.2))*(B03_3-1)/254)
L441{dom=value.dom;vr=-200:10000:0.001}:-5.1+((241.1-(-5.1))*(B03_4-1)/254)
L541{dom=value.dom;vr=-200:10000:0.001}:-1.0+((31.06-(-1.0))*(B03_5-1)/254)
L6H41{dom=value.dom;vr=-200:10000:0.001}:=3.2+((12.65-(3.2))*(b6_1-1)/254)
L6L41{dom=value.dom;vr=-200:10000:0.001}:=0.0+((17.04-(0.0))*(b6_2-1)/254)
L741{dom=value.dom;vr=-200:10000:0.001}:-0.35+((10.80-(-0.35))*(B03_6-
1)/254)
// REFLECTANCE PER TM //
//reflectance = pi*L*d^2/(Esun*coszeni)
R141{dom=value.dom;vr=0:1.00:0.001}:=pi*L141*(d(41))^2/(1969*coszen)
R241{dom=value.dom;vr=0:1.00:0.001}:=pi*L241*(d(41))^2/(1840*coszen)
R341{dom=value.dom;vr=0:1.00:0.001}:=pi*L341*(d(41))^2/(1551*coszen)
R441{dom=value.dom;vr=0:1.00:0.001}:=pi*L441*(d(41))^2/(1044*coszen)
R541{dom=value.dom;vr=0:1.00:0.001}:=pi*L541*(d(41))^2/(225.7*coszen)
R741{dom=value.dom;vr=0:1.00:0.001}:=pi*L741*(d(41))^2/(82.07*coszen)
// PLANETARY ALBEDO (rp) //
//The empirical equation used is as given by Liang//
rp{dom=value.dom;vr=0:1.0:0.001}:=0.356*R141+0.13*R341+0.373*R441+0.085*
R541+0.072*R741-0.0018
// Calculate broadTM surface albedo//
r0{dom=value.dom;vr=0:1.00:0.001}=(rp-0.03)/(0.78^2)
// NDVI, SAVI and LAI//
ndvi{dom=value.dom;vr=0:1:0.001}:=((R441-R341)/(R441+R341))
WNDVI =R441-1.387*R341
Lfactor=1-2*1.6*NDVI*WNDVI
SAVI=(R441-R341)*(1+Lfactor)/(R441+R341+Lfactor)

```

$$LAI=(SAVI-0.13)/0.47$$

// Surface temperature//

$$\text{BrightTemp} = -146.691 + 0.535 * (1282.71 / \ln((666.09 / L6H41) + 1)) + 0.897 * (1282.71 / \ln((666.09 / L6L41) + 1))$$

$$\text{surfem} = \text{iff}(\text{NDVI} > 0.16, 1.009 + 0.047 * \ln(\text{NDVI}), \text{iff}(\text{NDVI} < -0.1, 1.092))$$

$$\text{surfemis} = \text{iff}(\text{surfem} > 1.0988, \text{surfem})$$

$$\text{surfTemp}\{\text{DOM}=\text{VALUE.DOM}; \text{VR}=250:400:0.001\} = \text{BrightTemp} * (\text{surfemis})^{-0.25}$$

## A.2 Statistical Analysis

### A.2.1 Measured Data Entry Field 34: ANOVA

Data for Group A

$$A01 = \sqrt{120} \quad A02 = \sqrt{116} \quad A03 = \sqrt{133} \quad A04 = \sqrt{147} \quad A05 = \sqrt{\quad}$$

Data for Group B

$$B01 = \sqrt{140} \quad B02 = \sqrt{130} \quad B03 = \sqrt{120} \quad B04 = \sqrt{116} \quad B05 = \sqrt{\quad}$$

Data for Group C

$$C01 = \sqrt{196} \quad C02 = \sqrt{148} \quad C03 = \sqrt{164} \quad C04 = \sqrt{203} \quad C05 = \sqrt{\quad}$$

Data for Group D

$$D01 = \sqrt{200} \quad D02 = \sqrt{187} \quad D03 = \sqrt{200} \quad D04 = \sqrt{197} \quad D05 = \sqrt{\quad}$$

### A.2.2 ANOVA: Results

The results of a ANOVA statistical test performed at 10:03 on 22-OCT-2006

Source of Variation	Sum of Squares	d.f.	Mean Squares	F
Between	1.4662E+04	3	4887.	18.94
Error	3096.	12	258.0	
Total	1.7757E+04	15		

The probability of this result, assuming the null hypothesis, is 0.000

---

Group A: Number of items= 4

116. 120. 133. 147.

Mean = 129.

95% confidence interval for Mean: 111.5 thru 146.5

Standard Deviation = 14.0

Hi = 147. Low = 116.

Median = 126.

Average Absolute Deviation from Median = 11.0

---

Group B: Number of items= 4

116. 120. 130. 140.

Mean = 126.

95% confidence interval for Mean: 109.0 thru 144.0

Standard Deviation = 10.8

Hi = 140. Low = 116.

Median = 125.

Average Absolute Deviation from Median = 8.50

---

Group C: Number of items= 4

148. 164. 196. 203.

Mean = 178.

95% confidence interval for Mean: 160.3 thru 195.2

Standard Deviation = 26.1

Hi = 203. Low = 148.

Median = 180.

Average Absolute Deviation from Median = 21.8

Group D: Number of items= 4

187. 197. 200. 200.

Mean = 196.

95% confidence interval for Mean: 178.5 thru 213.5

Standard Deviation = 6.16

Hi = 200. Low = 187.

Median = 198.

Average Absolute Deviation from Median = 4.00

### A.2.3 Estimated Data Entry Field 35: ANOVA

Data for Group A

A01= A02= A03= A04= A05=

Data for Group B

B01= B02= B03= B04= B05=

Data for Group C

C01= C02= C03= C04= C05=

Data for Group D

D01= D02= D03= D04= D05=



#### A.2.4 ANOVA: Results

The results of a ANOVA statistical test performed at 10:12 on 22-OCT-2006

Source of Variation	Sum of Squares	d.f.	Mean Squares	F
Between	1.2881E+04	3	4294.	2061.
Error	25.00	12	2.083	
Total	1.2906E+04	15		

The probability of this result, assuming the null hypothesis, is 0.000

---

Group A: Number of items= 4

132. 132. 135. 135.

Mean = 134.

95% confidence interval for Mean: 131.9 thru 135.1

Standard Deviation = 1.73

Hi = 135. Low = 132.

Median = 134.

Average Absolute Deviation from Median = 1.50

---

Group B: Number of items= 4

130. 130. 130. 130.

Mean = 130.

95% confidence interval for Mean: 128.4 thru 131.6

Standard Deviation = 0.00

Hi = 130. Low = 130.

Median = 130.

Average Absolute Deviation from Median = 0.00

---

Group C: Number of items= 4

183. 183. 183. 183.

Mean = 183.

95% confidence interval for Mean: 181.4 thru 184.6

Standard Deviation = 0.00

Hi = 183. Low = 183.

Median = 183.

Average Absolute Deviation from Median = 0.00

---

Group D: Number of items= 4

191. 191. 195. 195.

Mean = 193.

95% confidence interval for Mean: 191.4 thru 194.6

Standard Deviation = 2.31

Hi = 195. Low = 191.

Median = 193.

Average Absolute Deviation from Median = 2.00