STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN (Glycine max. L. MERRILL) SEED ACROSS VARYING ENVIRONMENTS

MUDENDA HAMPANGO

A Dissertation Submitted to The University of Zambia in Partial Fulfillment of The Requirements for The Degree of The Master of Science in Plant Breeding and Seed Systems

SCHOOL OF AGRICULTURAL SCIENCES, DEPARTMENT OF PLANT SCIENCES

UNIVERSITY OF ZAMBIA

LUSAKA

2016

DECLARATION

I, Mudenda Hampango hereby declare that the work presented in this dissertation is my own and has never been submitted for a degree at this or any other university

Signature.....

Date.....

APPROVAL

This dissertation of Ms. Mudenda Hampango was approved as fulfilling part of the requirements of the award of Master of Science in Plant Breeding and Seed Systems by the University of Zambia

Examiner's Name and signature	Date
••••••	
•••••	

ABSTRACT

Soybean is an important crop with many important properties. The use of soybean varies from stock feed, biodiesel, edible oils and high protein food products, to soil amendment resulting from its nitrogen fixing ability and is a profitable cash crop. Zambia has a number of soybean varieties which have been developed with little consideration to adaptation for oil and protein to varying environments. Based on this, the stability of protein and oil content of soybean was evaluated across the three agro ecological zones of Zambia. The main objective of the study was to identify environments for high soybean oil and protein content production. Multi-location field trials involving fifteen soybean genotypes obtained from IITA, ZARI and SeedCo were conducted during the 2013/2014 rain season at five locations across the three agro ecological zones of Zambia namely; Masumba in Region I, Msekera, Kabwe and Golden Valley Agriculture Research Trust (GART) in Region II and Misamfu in Region III. A randomized complete block design with three replications was used at each site and the oil and protein content was determined using Near Infrared Reflectance. Analysis of variance (ANOVA) was used to determine statistical differences in performance of the genotypes for the studied traits while Additive Main Effect and Multiplicative Interaction (AMMI) Model was employed as a stability analysis tool. Results showed that both oil and protein content were significantly (p<0.05) affected by environments, genotypes and genotype by environment interactions, indicating differences in locations, presence of genetic variability among genotypes as well as the differential response of genotypes to environments. The mean oil content for the fifteen soybean genotypes ranged from 16.73% to 19.47% while the mean protein content ranged from 33.09% to 37.57%. In terms of locations, the highest mean oil content was obtained at Msekera (18.98%), while the lowest mean was from GART (16.38%). For protein content, the highest location mean was obtained from GART (38.23%) while the lowest was from Misamfu (33.47%). Msekera and GART can therefore be recommended for screening and production of genotypes with high oil and protein content respectively. AMMI indicated that the genotypes Lukanga, Safari, TGx 1988-22F and TGx 1740-2F were best suited for Msekera for high oil content while genotypes TGx 1830-20E, TGx 1987-23F, TGx 1887-65F and TGx 1988-22F were best suited for obtaining high protein content at GART. AMMI further indicated that TGx 1989-60F was the most stable genotype for oil content while for protein content, the genotypes TGx 1740-2F, Magoye and TGx 1988-18F were the most stable. These genotypes can be recommended for use as parental lines for developing soybean varieties that are stable for oil and protein content respectively.

DEDICATION

To my adorable children, Chikomo Jamel and Luumuno Elizabeth, and my beloved husband, Chikomo Memory Sikwangala.

ACKNOWLEDGEMENT

I wish to extend my sincere gratitude to my supervisors Dr. M. Mwala, Dr. D. Lungu and Dr. H. Agrama for their guidance and support during my research. I acknowledge their invaluable contribution through their comments and corrections during the research and dissertation writing. Special thanks go to the International Institute of Tropical Agriculture (IITA) for sponsoring my studies and the Department of Crop Sciences at University of Illinois for carrying the soybean oil and protein content Near Infrared Reflectance (NIR) analysis. I am also grateful to UNZA School of Agriculture soil science laboratory for carrying laboratory analysis on soil samples.

I would also like to thank the members of staff at IITA particularly those who worked under the SIMLEZA project for their assistance during the field trials.

I am also grateful to my employers, the Seed Control and Certification (SCCI) for securing my study leave and thereby allowing time to pursue my graduate studies. I also wish to thank my colleagues Mr. R. Chanda, Mr. S. Mwamba and Mr. I. Sichilima for their help in analysis.

Sincere thanks to my beloved husband and children for their love, support and patience during my studies. I also thank my parents and siblings for their encouragement, prayers and support throughout my studies.

DECLARATION	I
APPROVAL	II
ABSTRACT	111
DEDICATION	IV
ACKNOWLEDGEMENT	v
TABLE OF CONTENTS	VI
LIST OF FIGURES	XI
LIST OF APPENDICES	XII
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 LITERATURE REVIEW	6
2.0 SOYBEAN CROP	6
2.1 ORIGIN AND DISTRIBUTION OF SOYBEAN	6
2.2 CLASSIFICATION	6
2.3 BOTANICAL AND MORPHOLOGICAL CHARACTERISTICS	7
2.4 CHEMICAL COMPOSITION, USES AND IMPORTANCE	8
2.5 PRODUCTION, DEMAND AND CONSTRAINTS	9
2.6 GROWING CONDITIONS FOR SOYBEAN	11
2.6.1 TEMPERATURE AND PHOTOPERIOD	11
2.6.2 RAINFALL	12
2.6.3 SOIL	12
2.8 FACTORS AFFECTING VARIATION OF SOYBEAN OIL AND PROTEIN CONTENT	13
2.10 GENOTYPE BY ENVIRONMENT INTERACTION	15
2.11 ADAPTABILITY VS. STABILITY	17
2.12 STABILITY ANALYSIS	
2.13 AMMI	19
2.13.1 AMMI BIPLOTS	
2.13.2 AMMI STABILITY VALUE	21
2.14 NEAR INFRARED REFLECTANCE SPECTROSCOPY	22

TABLE OF CONTENTS

2.15	PLANT PROTEIN COMPOSITION	22
2.16	SOME IMPORTANT ELEMENTS IN AMINO ACID AND PROTEIN FORMATION	23
2.16.1	1 SULPHUR	23
2.16.2	2 IRON	23
2.16.3	3 Phosphorus	23
CHA	PTER 3 MATERIALS AND METHODS	25
3.1	TRIAL LOCATIONS	25
3.2	GENOTYPES USED IN THE STUDY	28
3.3	EXPERIMENTAL DESIGN AND MANAGEMENT	30
3.3.1	FIELD LAYOUT AND DESIGN	
3.3.2	TRIAL PLANTING AND CULTURAL PRACTICES	
3.4	DATA COLLECTION	30
3.5	STATISTICAL ANALYSIS	31
3.5.1	ANALYSIS OF VARIANCE	
3.5.2	CORRELATION ANALYSIS	
3.5.3	MULTIPLE LINEAR REGRESSION	
3.5.2	STABILITY ANALYSIS	
CHA	PTER 4 RESULTS	35
4.1	SOYBEAN OIL AND PROTEIN CONTENT	35
4.1.1	OIL CONTENT	
4.1.2	PROTEIN CONTENT	20
4.3		
	STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN	45
4.3.1	STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT	45
4.3.1 4.3.1.	STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT .1 THE AMMI STABILITY VALUE FOR OIL CONTENT	45 45 52
4.3.1 4.3.1. 4.3.2	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT .1 THE AMMI STABILITY VALUE FOR OIL CONTENT AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN PROTEIN CONTENT 	45 45 52 53
4.3.1 4.3.1. 4.3.2 4.3.2.	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT .1 THE AMMI STABILITY VALUE FOR OIL CONTENT AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN PROTEIN CONTENT .1 THE AMMI STABILITY VALUE FOR PROTEIN CONTENT	45 45 52 53 60
4.3.1 4.3.1. 4.3.2 4.3.2.	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT .1 THE AMMI STABILITY VALUE FOR OIL CONTENT AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN PROTEIN CONTENT .1 THE AMMI STABILITY VALUE FOR PROTEIN CONTENT .1 THE AMMI STABILITY VALUE FOR PROTEIN CONTENT	45 45
4.3.1 4.3.1. 4.3.2 4.3.2. CHA	STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT 1 THE AMMI STABILITY VALUE FOR OIL CONTENT AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN PROTEIN CONTENT 1 THE AMMI STABILITY VALUE FOR PROTEIN CONTENT PTER 5 DISCUSSION	45 45 52 53 60
4.3.1 4.3.1. 4.3.2 4.3.2 <u>CHA</u>	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT .1 THE AMMI STABILITY VALUE FOR OIL CONTENT AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN PROTEIN CONTENT .1 THE AMMI STABILITY VALUE FOR PROTEIN CONTENT	45 45 52 53 60 61
4.3.1 4.3.1. 4.3.2 4.3.2 <u>CHA</u> 5.1	STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN OIL CONTENT 1 THE AMMI STABILITY VALUE FOR OIL CONTENT AMMI MODEL AND PATTERN ANALYSIS FOR SOYBEAN PROTEIN CONTENT 1 THE AMMI STABILITY VALUE FOR PROTEIN CONTENT PTER 5 DISCUSSION OIL AND PROTEIN CONTENT VARIATIONS OF SOYBEAN GENOTYPES	45 52 53 60 61
4.3.1 4.3.1. 4.3.2 4.3.2 <u>CHA</u> 5.1 5.1.1	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN	45 45 52 53 60 61 61
4.3.1 4.3.1. 4.3.2 4.3.2 4.3.2. <u>CHA</u> 5.1 5.1.1 5.1.2	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN	45 45 52 60 61 61 61
4.3.1 4.3.1. 4.3.2 4.3.2 5.1 5.1.1 5.1.2 5.2	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN	
4.3.1 4.3.1. 4.3.2 4.3.2 4.3.2. 5.1 5.1 5.1.1 5.1.2 5.2	 STABILITY ANALYSIS FOR OIL AND PROTEIN CONTENT IN SOYBEAN	

CHAPTER 7	REFERENCES	76
CHAPTER 8	APPENDICES	94

LIST OF TABLES

Table 1a: General Features of the Trial Sites with soil and climatic data during the 2013/ 2014 season
Table 1b: Monthly Meteorological data of sites used in the study during the 2013/2014 season
Table 1c: Soil Analysis Results for sites used in the study
Table 2: Genotypes used in the Study
Table 3: Planting Dates for the five Locations
Table 4. Analysis of variance per location for soybean oil content of 15 soybean genotypes
Table 5: Mean Oil content (%) of fifteen soybean genotypes tested at the five different locationsduring the 2013/ 2014 season
Table 6: Combined analysis of variance across five locations for soybean oil content
Table 8: Mean Protein content (%) of fifteen soybean genotypes tested at the five differentlocations during the 2013/2014 season
Table 9: Combined analysis of variance across five locations for soybean protein content
Table 11. Mean Oil content and ranking order with IPCA 1 and IPCA 2 scores of fifteen genotypes tested across five environments
Table 12: Environment means and IPCA scores of the fifteen genotypes in the five environments 49
Table 13: The AMMI model's first four genotype selections for mean soybean oil content across five environments. 51

Table 15: ANOVA for the AMMI Analysis of soybean protein content across five
environments
Table 16. Mean protein content and ranking order with IPCA 1 and IPCA 2 scores of fifteen
genotypes tested across five environments
Table 17. Environmental means and IPCA scores of the five environments for protein content
(%)
Table 18: The AMMI model's first four genotype selections for mean soybean protein content
across five environments
Table 19. AMMI Stability Value (ASV) and ranking order with IPCA 1 and IPCA 2 scores for
soybean Protein content of fifteen genotypes tested across five environments

LIST OF FIGURES

Figure 1: Map of Zambia showing the five experimental sites	
Figure 2: Pedigree of IITA genotypes used in the study	29
Figure 3: AMMI 1 biplot of genotypes and environment IPCA 1 scores versus the so	ybean oil
means of fifteen genotypes and five environments	
Figure 4: AMMI 2 Biplot of IPCA1 scores versus the IPCA2 scores of fifteen genotypes	and five
environments for soybean oil content	50
Figure 5. AMMI 1 Biplot of genotypes and environment IPCA1 scores versus the soybea	n protein
means of fifteen genotypes and five environments	54
Figure 6. AMMI 2 Biplot of IPCA1 scores versus the IPCA2 scores of fifteen genotypes	and five
environments for soybean protein content	58

LIST OF APPENDICES

Appendix 1: Correlation of oil and climatic parameters	94
Appendix 2: Correlation of protein and climatic parameters	95
Appendix 3: Stepwise Regression: Oil versus Rainfall, Temperature and Edaphic factors	96
Appendix 4: Stepwise Regression: Protein versus Rainfall, Temperature and Edaphic	97

CHAPTER 1 INTRODUCTION

Soybean (*Glycine max* (L) Merrill) belongs to the family Leguminosae and genus *Glycine* L. It is believed to have been first domesticated around the eleventh century in China (Lance and Garren, 2005).

The crop is an important legume with multifarious uses and its cost effectiveness is ensured through biological nitrogen fixation and rotation with exhaustive crops since it replenishes and maintains soil fertility (Ngalamu, 2013). Soybean has a wide adaptation being cultivated in tropical, subtropical, and temperate climates.

Soybean has been grown for about 3,000 years in Asia and more recently, has been successfully cultivated in different parts of the world (soyatech, 2012). The crop has evolved to be an important crop at global level from the time of its domestication till now. Today, the world's top producers of soybean are the United States, Brazil, Argentina, China and India. Its demand is dominated by the USA, China and Europe (FAOSTAT, 2012). About 85 percent of the world's soybeans are crushed or processed annually into soybean oil and meal. Approximately 98 percent of the crushed soybean meal is further processed into animal feed and the remainder is used to make soy flour and proteins. On the other hand, 95 percent of the oil is consumed as edible oil while the balance is used to make industrial products such as fatty acids, soaps and biodiesel. Soybean contains all the eight amino acids that are essential for human health making it one of the few plants that provides a complete protein. (Soyatech, 2012). Some nutritional advantages could therefore be obtained by replacing many animal based foods for soybean foods because soybean represents an excellent source of high quality protein with a low content in saturated fat and a great amount of dietary fibre. Owing to this, the possible use of soybean in functional food design has attracted a lot of interest, since the consumption of soybean protein and dietary fibre is claimed to reduce the risk of cardiovascular diseases as well as improving glycemic control. Furthermore, soybean isoflavones are associated with a potential role in the prevention and treatment of different diseases, soybean therefore could play an important role in the promotion of health (Aparicio et al., 2008).

Despite the clear benefits of soybean, small holder production in Zambia remains low with the smallholder farmers meeting only 15% of local demand, while remaining 85% is supplied by commercial farmers (Technoserve, 2011). Commercial yields are much higher than smallholder yields as commercial farmers tend to employ better agronomic practices. Soybean seeds are

composed of approximately 20% oil, 40% protein, 30% carbohydrate, 9% crude fiber, and 5% ash (Padgette, 1996). Soybean seed is known to have the highest protein content among all food crops and is second only to groundnut in terms of oil content among food legumes (Mpepereki, 2001). Soybean oil contains approximately 12% palmitic acid (16:0), 4% stearic acid (18: 0), 23% oleic acid (18:1), 53% linoleic acid (18:2), and 8% linolenic acid (18:3) (Lee *et al.*, 2009). The soybean protein is rich in lysine, tryptophan, threonine, isoleucine and valine and therefore complements well with cereal grains that are deficient in those amino acids (Kumar *et al.*, 2010).

Soybean was introduced in Zambia in the 1930s, but remained a minor crop grown mostly by the commercial farmers. The crop is however currently grown by both small and large scale farmers as it has gained popularity owing to its industrial properties and nutritional benefits. Soybeans are adapted to regions II and III of Zambia and will grow well wherever maize grows (Miti, 1997). However, the varying agro-ecological conditions found in these regions mean that it is extremely important to select appropriate cultivars for each environment.

The Zambian soybean market has attained self-sufficiency and grown rapidly with some exports. The production is dominated by commercial farmers and there is considerable scope for production growth. With production of 112,000 MT and consumption of 90,000 MT in 2009/10, Zambia is a net exporter of soybean, and has been so in the recent past. Despite the volatility of the soybean market in Zambia, production has grown over the years and this has been largely achieved by increasing planting area rather than increasing yields (Technoserve, 2011).

In Zambia, soybean is mostly used in production of edible oils and the cake is made into by products such as soya chunks, High Energy Protein Supplements (HEPS) and soya meal for human consumption, in addition to being a valuable stock feed. Soybean has also been used successfully as a constituent for making bread and cakes in Zambia. The growing demand of soybean offers significant opportunity for smallholder farmers to improve their cash base (Lubungu *et al.* 2013).

Development of soybean varieties with superior chemical compositions that meet special food requirements and applications is now a major research priority. During selection of soybeans for a particular seed breeding program or particular food application, it is important to know the major factors that affecting soybean quality with regards to protein and oil contents (Liu *et al.*, 1995). The nutrient composition of soybean seed can be influenced by several factors such as the genetics, cultivar, growth conditions, as well as processing. Therefore, significant variations can be detected

in the soybeans and soybean based foods nutritional profile (Grieshop and Fahey, 2001). Factors that directly affect Soybean plants include high temperature, day length and relative humidity, and these may cause the potential yields not to be realized in all environments (Uncu and Arioglu, 2005).

The versatile nature of this crop and its increasing contribution to industrial, agricultural and medicinal sectors call for concerted pre-breeding efforts to widen the genetic base and for formulating an effective breeding programme to develop varieties suitable to specific agro climatic conditions (Dayaman, 2007) for various traits of interest.

In the production of soybeans, performance and, therefore, expression of plant characters is a manifestation of genotype by environment interactions (GE). Genotype by environment interactions can simply be defined as the combined effect of the genetic structure of a plant and environmental factors in the final performance of the organism. Genotypes by environment interactions are ascribed to differences in sensitivity, which means that a given environment affects some genotypes more than others (Falconer, 1981). Genotypes by environment interactions represent differential responses of genotypes and renders mean performance less useful as genotypes' relative ranking or degree of magnitudes vary across the environments (Allard and Bradshaw, 1964).

Knowledge of the pattern and magnitude of genotype by environment interactions and stability analysis is important for understanding the response of different genotypes to varying environments and for identification of widely adapted and specifically adapted genotypes. This is crucial in plant breeding in order to rationalize resources and confine genotype testing to sites with informative data facilitating a rapid response to selection. (Tukamuhabwa *et al.*, 2012)

Stability analysis is an important and efficient tool for plant breeders and agronomists. It helps to identify and select the most stable, high performing genotypes that are best suitable under a given set of environmental conditions (Jandong *et al.*, 2011).

The prevailing phenomenon of global climate change may result in strong impacts on agriculture, especially on crop growth and yield. Crop performance is largely determined by climatic conditions during the growing season, as such, even minor deviations from optimal conditions can seriously threaten the quality of harvest. Knowledge on how environmental factors affect crop growth and development could therefore help in reducing the possibilities of significant quality and yield loss

and in turn improve the selection of specific cultivars to be grown in target regions (Marjanovic-Jeromela *et al.*, 2011)

Genotype by environment interactions have effects on biochemical and physical characteristics of soybean seed (Kumar *et al.*, 2006) and these have been reported to affect soybean protein and oil content (Piper and Boote, 1999).

There is need to characterize soybeans adaptation with regards to oil and protein content under different agro-climatic conditions of Zambia. Although many soybean varieties are recommended for cultivation in Zambia, the information on their stability with respect to oil and protein content is lacking. Given the ongoing practice of crop diversification and adoption of maize- legume mixed systems in Zambia, soybeans could be distributed across the country expecting stable performance to be economic and useful; this will not be realized with regards to grain yield, protein and oil content unless these traits were assessed for their stability across varying environments. An effort in this direction is cardinal in order to popularize the crop. The challenging impacts of climate change make this effort more urgent.

Wide adaptation to particular environments and consistent performance of recommended varieties/genotypes are very cardinal for successful cultivation of soybean. Although many soybean varieties are recommended for cultivation in Zambia, the information on their stability with respect to oil and protein content is lacking.

In the present study, a number of advanced soybean genotypes coming out of the International Institute of Tropical Agriculture (IITA) breeding programme as well as some released varieties were evaluated at several locations in order to identify environments for high soybean oil and protein content.

The specific objectives were:

- i. To determine performance of soybean genotypes with regard to oil content.
- ii. To determine performance of soybean genotypes with regard to protein content.
- iii. To characterize soybean genotypes for oil content across selected environments in Zambia.

iv. To characterize soybean genotypes for protein content across selected environments in Zambia.

The underlying assumption/hypothesis is that there is sufficient genetic variation in soybean genotypes to identify those that would perform well with respect to oil and protein content, in specific environments or perform well across varying environments.

Information from this study will be useful in guiding soybean production programmes in targeted environments by having the crop grown where it is most suited with respect to soybean oil or protein content thereby increasing production of quality crop. The information will also be useful in the selection of stable lines with regard to the studied traits to be used as parental crossing lines in breeding programmes therefore enhancing the efficiency of breeding and selection for superior soybean varieties and improving production of quality crop. It is envisaged that the results will also benefit plant breeders, agronomists and consumers as the research will also provide information on soybean quality traits and their stability that has not previously been available in the country.

CHAPTER 2 LITERATURE REVIEW

2.0 Soybean Crop

Soybean *Glycine max* (L.) Merrill is an annual food crop native to East Asia, predominantly China. Soybean is classified as an oil seed rather than a pulse and is very important in both human and livestock nutrition.

2.1 Origin and Distribution of soybean

Soybean is considered as one of the oldest cultivated crops. While it is difficult to pinpoint when the domestication of soybean first began, Hymowitz and Newell (1981) indicate that the first written record of soybean appeared in the Book of Odes, written during the 11th to 7th centuries B.C.

Scholars generally agree that the cultivated soybean (*Glycine max*) originated from North China in the eleventh century B. C. or perhaps a bit earlier (Fukuda, 1933 and Singh, 2010), from where it spread to other parts of the world. From Asia, the crop was introduced into Europe, America and later to Africa by Chinese traders along the east coast of Africa in the early 19th century and is now widespread across the continent. Soybean is widely grown on large scale in both the temperate and tropical regions such as China, Thailand, Indonesia, Brazil, the USA and Japan where it has become a major agricultural crop and a significant export commodity (Evans, 1996).

2.2 Classification

Soybean belongs to the family fabaceae also known as leguminosae (pea family), sub family faboideae, genus *glycine*, and species - *max* [(L). Merrill]. Its bionomial name is therefore *Glycine max* (L.) Merr.

The cultivated soybean is a diploidized tetraploid (2n=40) and was first described in 1753 by Linnaeus as both *Phaseolus max*, based on the specimen he saw, and Dolichos soja, which he compiled from the descriptions of other writers (Hymowitz *et al.*, 1981) The combination *Glycine max* (L.) Merr. as proposed by Merrill in 1917 has since become the valid name for this important crop.

The genus Glycine is composed of two subgenera, Glycine (perennials) and Soja (annuals). It is believed that the cultivated soybean was domesticated from the annual wild soybean *Glycine soja*

Sieb.*et* Zucc. which grows wild in Taiwan, Korea, Japan and China. *G. gracilis*, an intermediate between *G. soja* and *G. max*, is said to have been observed in Northeast China (Skvortzow, 1927).

There are some significant differences between *Glycine max* and *Glycine soja* at the genomic level, which corresponds with some functionally important genes for seed, oil and protein traits. These genes can explain some of the phenotypic differences observed between *G. soja* and *G. max* especially in terms of the above three traits. The Seed oil concentration of *G. soja*, is about 8% while that of *G. max* is about 25%. (Joshi *et al.*, 2013). Highly productive and high protein lines have been derived from soybean and G. *soja* hybrids (Hartwig, 1973). With its high protein content, the wild soybean is also used as fodder in some regions of China. (Hong and Blackmore, 2015)

2.3 Botanical and Morphological Characteristics

Soybean is an annual, erect hairy bushy herbaceous plant, ranging in height between 30 and 183 cm, depending on the genotype (Carlson 1973; Ngeze, 1993). It is an erect bushy herb with twinning aerial weak stem and has a taproot system along with large number of fibrous, secondary roots and root nodules. It is has determinate as well as indeterminate growth habit (Martin, 1984).

The plants have a tap-root, which can be as long as 2 m, with numerous lateral roots. Inflorescence on each plant comprises one or two self-fertile flowers that are borne in the axils of the leaves. The colour of the flower varies depending on the cultivar with white and purple represented. The flowers also have tiny hairs (pubescence) that are either or grey or tawny coloured. The fruit is a straight or slightly curved pod that grows in clusters of three to five and varies in length from two to seven centimeters. The pod comprises of two halves of a single carpel joined together by a dorsal and ventral suture. One plant can produce up to 400 pods, with about two to twenty pods at a single node (OECD, 2000).

The pods usually contain two to four seeds whose size could be five to seven millimetres in diameter. The shape of the seed is usually oval but tends to vary among cultivars from almost spherical to elongate and flattened (Hymowitz, 1995). Seeds tend to be yellow in colour with either a dull or a shiny seed coat and the colour of the hilum ranges from yellow to black, with black being most common (Koivisto, 2003;Acquaah, 2007) with yellow or green cotyledons.

2.4 Chemical Composition, Uses and Importance

Soybean seed is composed of an average of 40% protein, 30% carbohydrate, 20% oil and 5% ash on a dry weight basis (Padgette, 1996; SoyStats, 2011;). It is the world's leading source of oil and protein. It has the highest protein content among the cultivated food crops in the world and is second only to groundnut in terms of oil content among food legumes (Mpepereki, 2001; Gurmu *et al*, 2009). Most crop plants are specialized to produce either high protein contents or large amounts of energy in the form of sugars and oils. Soybean on the other hand does both, by moving large proportions of its stored nitrogen and energy from all parts of the plant into the seeds as they mature. This mobilization of reserves from the stems, leaves, and roots into the seeds is so extreme to the extent that the plant is unable to survive and dies within a few weeks, but not before producing large amounts of protein and oil-rich seeds (Boucher, 2011)). With approximately 40 percent protein and 20 percent vegetable oil by weight in its seeds, soybean stands out as an extraordinary source of both protein and energy (Wiggins, 2012). However, literature indicates that there is a limitation in utilization of soybean for human consumption as it consists of some anti-nutritional factors (trypsin inhibitors) which require soybean to be treated or processed before consumption (Lokuruka, 2010).

The use of soybean ranges from feed, biodiesel, edible oils and other food products. Soybean is a complete protein and soy- based foods are rich in vitamins and minerals. Soybean meal is a valuable and desirable product as soybean protein provides all the essential amino acids in the amounts needed for human health (El- Shemy, 2011; Soyatech, 2012) and livestock feeding (Carrera *et al.* 2011). The rich protein content in soybeans also means it could contribute to improved nutritional status of rural households and can be an excellent substitute for meat for resource poor families that cannot afford protein rich foods such as meat, fish, eggs and milk which are often scarce and expensive.

Soybean flour is most widely used in baked goods; 2%-15% is added to breads, crackers, muffins, donuts, cakes, rolls, cookies, tortillas, or chapatis. It is also used in pasta products (spaghetti, noodles, macaroni), processed meats (sausages, bologna, frankfurters, meat loaves), gravies, sauces, soups, cereals, prepared mixes (pancake and waffle), dairy substitutes, candies (caramels and toffees)special diet foods (diabetic, allergenic, high protein), and spice bases (Shurtleff and Aoyagi, 2004).

Soybean oil on the other hand is the world's most widely used edible oil as it has no cholesterol as this is inherently absent in most plant foods (USDA). Soybean oil has a natural taste and nearly

imperceptible odour, which makes it the ultimate choice of vegetable oil for both domestic and industrial food processing (Mpepereki *et al.*, 2000; Mahama, 2011).

Other notable attributes of soybean oil include its polyunsaturated fatty acids, biodiesel properties and usefulness in food products as both oil and as the emulsifier, lecithin (Carpenter *et al.*, 2002). Soybean oil is also used as an ingredient in other foods and manufactured products such as paints, inks, dyes, biodiesel and lubricants (Aquaah, 2005). And it is also heavily used in industries, especially in the manufacture of soap, plastic products, glycerine and enamels (Ngeze, 1993; Rienke and Joke, 2005; Mahama, 2011).

A major food use of soybean in developed countries is as purified oil, utilised in margarines, shortenings and salad oils. It is also used in different food products, such as tofu, soya sauce, simulated milk and meat products. Soybean meal is used as a supplement in feed rations for livestock. Industrial uses of soybeans include production of yeasts and antibodies to the manufacture of soaps and disinfectants (FAO, 2014).

Soybean is an important component of most smallholder cropping systems in Africa. Its importance includes enhancing household food and nutrition security, as well as raising rural incomes and reducing poverty through provision of employment in soybean based activities.

Like other legumes, soybean also improves soil fertility by converting atmospheric nitrogen into ammonia and organic derivatives for its own use as well as for subsequent crops in rotation. It therefore cuts down on the amount of nitrogen fertilizer that farmers need to apply to their fields to improve productivity. This is a major benefit in Africa where soils are poor in nutrients and fertilizers are expensive and not readily available for farmers (IITA, 2009; Mahama, 2011).

2.5 **Production, Demand and Constraints**

Over the last 20-30 years, consistent improvements in average yield levels and reductions in production costs have steadily improved the competitive position of soybeans among arable crops. Among oil crops, soybean has prominent role at the global level. Today, soybean accounts for about 35% of the total harvested area dedicated to both annual and perennial oil crops. Soybean's share in global oilseed output is estimated at over 50 % (Thoenes, 2006).

The popularity and widespread cultivation of soybean can be attributed to its adaptability. The ability of soybean to grow in soils that are also suitable for growing maize, its level of drought tolerance, and its variety of non-food uses are all favorable qualities of this crop (Aquaah, 2005).

According to FAO (2012), total world production of soybean in 2010 was 261.6 million metric tonnes. The three major soybean producing countries in the world being USA (90.6 million metric tonnes), Brazil (68.5 million metric tonnes) and Argentina (52.6 million metric tonnes).

The cultivation of Soybean in Africa is negligible, with Africa contributing as little as less than 1% to global Soybean production (Swanby, 2010). Nigeria is the biggest producer of soya beans on the continent followed by South Africa.

However, the Soybean industry is well established in Southern Africa, with total production of 861,000 MT in 2010 and demand of 2 million MT. South Africa dominates both production and demand though Zambia, Zimbabwe and Malawi are also significant producers. Production is mostly by commercial farmers (who took up 84% of production in 2010); this however varies significantly depending on the country, with smallholder farmers dominating production in Mozambique and Malawi. Demand is dominated by Soybean cake for the poultry industry and Soybean oil for human consumption and is expected to continue to as high as 3.5 million MT by 2020 (Opperman and Varia, 2011).

Zambia has attained self-sufficiency in soyabeans production and the crushing capacity has almost doubled and is able to satisfy the local demand for soy cake. Historically, the market for soy in Zambia has been driven by the feed industry, particularly the poultry industry. However, there have been dramatic increases over the past 9-10 years, especially in demand for soya chunks and High Energy and protein Supplements for malnourished children (Technoserve, 2011).

Soybean is cultivated in nearly all the provinces of Zambia, though production levels vary and it is more adapted to region II and III. The national soybean production figures in recent past are as follows; 2008/2009 production was 119,000Mt, 2009/2010 production was 112,000Mt, while 2010/2011 production was 116,000Mt. This is against Soybeans national consumption of 136,000Mt.The total area planted to soya-beans for the 2010/2011 season however decreased by 1percent to 61,422 hectares from 62,331 hectares during the 2009/2010 season (CSO/MAL). In theory, there is no limit soybean production as Zambia has 33 million hectares available for

additional production. The Eastern province leads the country in small holder soybean production (Lubungu *et al*, 2013).

The cake market which averaged 90,000MT in 2009/ 2010 has been driven by the growth of the poultry industry which drives the demand for feed. This is expected to continue with the cake market's projected rising to 194,000 MT for domestic market by 2020, driven by a rise in demand from poultry from 65,000 MT in 2011 to 140 MT in 2020. The oil market is also large enough to absorb all of the soy oil produced in the country. This was estimated to be equivalent to 390,000 MT of soybeans in 2009/2010. Zambia is well placed to export soy to Zimbabwe, South Africa and the Democratic Republic of Congo, but high transport and inconsistent policy limit traded volumes (Technoserve, 2011).

Low yields averaging about less than 1 tonne per hectare in tropical Africa and a shortage of fertilizer constrain the ability of some countries to increase soybean production in Africa (IITA, website).

Small holder farmers in Zambia predominantly use recycled local soybean varieties. The reasons for predominance of local, recycled seed usage include; inexpensiveness of local self-pollinating varieties which can be recycled for more than five years with little reduction in production output. In addition to this, the, availability of commercially produced and open pollinated commercial seed varieties has been unreliable in the past.

Usage of most inputs in soybean production is very low among smallholders due to high cost, lack of availability and insufficient awareness of benefits. Many farmers purchase uncertified seed and no other inputs. For small holder farmers, lack of input accessibility is a big problem as many inputs are only sold in major markets. This is particularly true for certified seed, lime and inoculant. Many smallholders want to buy inputs at harvest time rather than planting time because that is when they have cash (Lubungu *et al.*, 2013).

2.6 Growing Conditions for Soybean

Soybean growth and development is influenced by factors such as;

2.6.1 Temperature and Photoperiod

Temperature plays a very important role in determining the rate of soybean growth. Soybean seeds require minimum soil temperature of about 15°C in order to germinate. The minimum temperature for soybean development is 10°C with the optimum being 22°C and maximum 40°C (Rienke and Joke, 2005).

Temperatures below 13°C hamper flowering in soybean. Different plant parts and growth stages however respond differently to the same temparture conditions (Magagane, 2012)

Soybean is a photo- period sensitive crop whose growth is influenced by daylight length. Vegetative growth before flowering begins is mainly affected by length of daylight. Soybean begins to flower as nights become longer. Different soybean cultivars however tend to have different daylight length requirements due to genetic variation which makes it possible to cultivate soybean in a variety of conditions (Youdeowei *et al.*, 1986).

2.6.2 Rainfall

Soybean requires optimum moisture in order for seed to germinate and grow well. Rainfall should be between 350mm and 750 mm, and well distribute throughout the growing season for soybean production. Water requirement is at its peak during the vegetative stage and decreases at maturity stage (Rienke and Jone, 2005; Mahama, 2011). Soybean also performs well in warm, dry areas under irrigation (Magagane, 2012). The crop is very sensitive excess as well as moisture stress during flowering and pod formation (Youdeowei *et al.*, 1986). Established soybean plants can withstand considerable drought (Martin, 1988).

2.6.3 Soil

The most ideal soils for soybean growth are loose, deep, well drained loamy textured soils varying from loamy sands to clay loams with a fine but firm seedbed and good water holding capacity. Soybeans can be grown in soils with a lower pH than other legumes as it is able to tolerate acidity. The crop does well when soil pH is between 5.5 and 7.0 (Ngeze, 1993) as these conditions enhance the availability of nutrients such as nitrogen and phosphorus as well as microbial breakdown of crop residues and symbiotic nitrogen fixation (Ferguson *et al.*, 2006; Mahama, 2011).

Soybean should not be planted on soils that are sandy, gravelly, or shallow in order to avoid drought stress. Waterlogged soils or soils with surfaces that can crust should also be avoided, as they can lead to poor seedling emergence (Dugje *et al.*, 2009) as a result of the hypocotyl breaking during emergence.

2.7 Importance of selection for oil and protein content in soybean

In soybean breeding, special attention is given to developing cultivars that have high protein and oil content, in addition to high and stable yields (Hollung *et al.*, 2005). Besides the interest in individual

soybean grain components, the processing industry equally finds the ratio between protein and oil content in soybean grain to be significant (Miladinovic *et al.*, 2011).

Historically, soybean breeders have mainly used total protein content and not oil content as a selection criterion for germplasm development. However, recently, both oil content and quality have drawn much attention in soybean genetics and breeding programs, due to the increased demand for vegetable oils and increased consumer awareness of health issues around dietary fats. A strong indirect phenotypic correlation exists between these traits. In addition, the variation in soybean germplasm for protein content is significantly higher than that observed for total oil content. (Clement and Cahoon, 2009).

Selection of soybean genotypes for soy-food offers potential for expanding the already growing international market. The increasing market for soy based foods and the health benefits associated with them indicate the economic potential and underscore the need to identify and develop soybean cultivars that are nutritional, high yielding and suitable for food processing and human consumption (Kuhn, 1996).

To satisfy the demand by producers and consumers, a number of soybean varieties with excellent seed quality and agronomic characteristics have been bred and released for cultivation by farmers in tropical Africa (FAO, 1999) with the aim of increasing production and enhancing protein intake of the low and middle income earners. The increasing interest from farmers has resulted in extension of soybean production to the high rainfall belts of sub-Saharan Africa (Mutsaers, 1991).

2.8 Factors affecting variation of soybean oil and protein content

Although genetics are generally considered the main determinant of composition (Thorne and Fehr, 1970; Wilcox, 1985; Burton, 1989; Bonato *et al.*, 2000), environmental variation is also a determinant of protein and oil concentration in soybean (Ojo *et al.*, 2002). Genotypic differences in soybean protein content are a result of additive gene action with heritability values ranging from medium to high (Jaureguy *et al.*, 2011; Rodrigues *et al.*, 2014). The differences between varieties generally represent fifty per cent of the total variation in the soybean seed composition (Brumm and Hurburgh, 2002).

The inverse proportional relation of oil and protein content is well known (Filho *et al.* 2001; Ojo *et al.*, 2002; Schwender *et al.*, 2003; Popovic *et al.*, 2013) and genetic and ecological factors influence the negative correlation of these two soybean seed constituents. On the other hand, high seed protein concentration is frequently associated with less yield (Carter *et al.*, 1982; Wilcox and Zhang, 1997; Wilcox and Shibles, 2001)

During soybean seed development the four main stages can be observed: morphogenesis and cell division, cell enlargement, seed maturation, and ultimately the release of moisture and period of seed dormancy. Synthesis of proteins and oils takes place during the growth phase of seed cells (Blanusa *et al.*, 2000). Therefore, the growing conditions at this stage are said to be significantly correlated with protein and oil content in soybean seed (Dordevic *et al.*, 2010). Several studies have reported that environmental conditions have the greatest effect on the oil and protein content of soybean seeds (Fehr *et al.*, 2003; Ning *et al.*, 2003; Zhang *et al.*, 2005). Numerous estimations, however, confound both genotypic and environmental effects, which makes it difficult to separate the relative importance of these two factors (Piper and Boote, 1999; Yaklich *et al.*, 2002; Dardanelli *et al.*, 2006).

In addition to genetics and environmental or growing conditions, the origin, soil characteristics, agronomic practices, technological processes (Piper and Boote, 1999; Intech, 2011) and water stress (Dornbos and Mullen, 1992; Noureldin *et al.*, 2002) have a bearing on the accumulation and concentration of chemical and mineral components in soybean seeds. Heidarzade et al., (2016), reported that climatic and edaphic factors affect soybean growth. Caliskan et al., (2008), also reported that micronutrients such as Iron (Fe), Manganese (Mn), Zinc (Zn) and Molybdenum (Mo) affect performance of soybean. Other researchers also reported that seed yield and protein content were affected more than oil content by environment conditions (Gurdeep-Sing *et al.*, 2001; Sudaric *et al.*, 2006; Arsnaloglu, 2011).

In the past, breeding programs primarily focused on increasing the yield of the crop grown under regional climatic conditions (Ustun *et al.*, 2001). For the processing industry however, the chemical composition of soybean seeds is one of the most important factors (Zilic *et al.*, 2009; Popovic *et al.*, 2013). Therefore, considering the demands of the processors in the recent times, soybean breeding has been focused on increasing the protein content, and improving oil quality (Miladinovic *et al.*, 2011; Tubic *et al.*, 2011; Popovic *et al.*, 2013).

2.9 Efforts made to deal with variation in soybean oil and protein content

High protein and low oil content add nutritional value to soy foods. Germplasm that cover a wide range in protein content (33.1 to 55.9%) and oil content (13.6 to 23.6%), are available for breeders in order to modify the seed protein/oil ratio. The negative correlation between protein and oil facilitates the development of high protein and low oil lines (Van Schoote, 2011).

Although there remains a strong economic incentive to develop cultivars with high protein and oil contents while maintaining a competitive yield, progress has been slow. Effective breeding techniques require accurate, inexpensive and reliable soybean composition analysis. Certain areas of breeding and selection research would also benefit from single soybean seed analysis (Silvela *et al.*, 1989). Conventional composition analysis methods such as the Kjeldahl method for protein measurement and the ether extraction method for oil fraction measurements are time consuming, expensive and impractical for measurements on large numbers of soybean samples required for molecular genetic mapping and other selection and breeding studies. In addition to problems such as low speed and high cost, wet-chemistry methods are destructive and rather inaccurate for single seed analysis, with the notable exception of the extracted protein determination by the Lowry method (1958) (Baianu *et al.*, 2010).

Most of the previous reports in literature have documented a highly negative correlation between soybean seed protein and oil, which would seem to be difficult to overcome when breeding for higher contents of these traits. However, in some reported populations, the correlation was observed to be slightly weaker, suggesting some progress could be made towards improving seed oil without simultaneously incurring a substantive decrease in seed protein. Pathan *et al* 2013, in their study of soybean constituents reported that heritability for soybean protein was high (0.76) while that of oil was 0.66 and these results were moderately similar to those reported by Panthee *et al.*, (2005). With such heritability estimates, soybean breeders might be able to use simple selection techniques with appropriate breeding methods to develop high oil breeding lines or cultivars (Sun, 2011).

2.10 Genotype by Environment Interaction

Genotypes tested in different locations or years often exhibit significant fluctuation in performance due to the response of genotypes to environmental factors such as rainfall, soil fertility, temperature, soil types or disease pathogens (Kang, 2004). These fluctuations are often referred to as genotype by environment interactions (GEI) and are common. GEI results in significant differences in the performance of genotypes when evaluated in different locations (Gauch and Zobel, 1997; Zhe *et al*, 2010).

This differential response of genotypes across environments (GEI) tends to limit response to selection and subsequently progress in plant breeding programmes (Crossa *et al.*, 1999; Alberts, 2004; Tukamuhabwa *et al.*, 2012). Understanding the cause of GEI is important as it helps in selecting varieties with the best adaptation and that can give stable yields. Varieties that show low GEI and have high stable performance are desirable for crop breeders and farmers, because that indicates that the environment has less effect on them and their good performance is largely due to their genetic composition. It is important to understand crop development in relation to biophysical conditions and changes in season when selecting well adapted genotypes (Linnemann *et al.*, 1995).

Multi-location trials are conducted for various agronomic and grain quality traits in order to identify superior genotypes across a wide range of environmental conditions. Beck *et al.* (1991) reported that when genotypes are grown under a wide range of environments and outside their usual adaptation zone, the occurrence of large GEI is expected. Large GEI makes it difficult for the identification of better performing genotypes. The GEI is of practical significance when the ranking of genotypes varies among environments; this is known as crossover interactions (Crossa and Cornelius, 1997; Russell *et al.*, 2003; Masindeni, 2013).

It is important for breeders to evaluate different types of cultivars under various environments for grain quality and other traits valuable to the end users. Significant GEI allows breeders to further assess the adaptability and overall stability of the genotypes across different environments. If GEI is well characterised, it is possible to develop locally adapted cultivars with high consistent performance.

Development of improved varieties of soybean, using exotic breeding materials, causes a change in photoperiodic response and general adaptation of the progenies. Therefore, to determine the pattern of genotype response to environment and prioritise genotypes for use in a breeding programmes, quantification of genotype by environment interactions is necessary (Gauch, 2006). This is important especially when dealing with advanced generation soybean lines not tested for adaptation to the main soybean producing areas of a given country (Tukamuhabwa *et al.*, 2012).

2.11 Adaptability vs. Stability

The terms 'stability' or adaptability refer to consistent high performance of genotypes across diverse sets of environments (Romagosa and Fox, 1993). Adaptability in a plant breeding context indicates the ability of a genotype to be high yielding with respect to a given environment or given conditions to which it is adapted (Gallais, 1992).

In other words, adaptability refers to the ability of genotypes to successfully assimilate environmental stimuli, which is advantageous from an agricultural yield standpoint; i.e., the adaptability is evaluated based on the average performance of the genotype (Da Silva, 2014).

When breeding for wide adaptation (i.e. adaptability), the aim is to obtain a variety which performs well in nearly all environments; while when breeding for specific adaptation, the aim is to obtain a variety which performs well in a definite subset of environments within a target region. The adaptive response of a variety is assessed with respect to other genotypes and tends to undergo modification when better performing germplasm becomes available (Annicchairico, 2002).

A stable genotype on the other hand according to Becker and Leon (1988), is one that possesses a constant performance regardless of any variation in environmental conditions (Fasahat *et al*, 2015). Therefore, a stable genotype can be referred to as one that is capable of utilising the resources available in high yielding environments and has a mean performance that is above average in all environments (Allard and Bradshaw, 1964; Eberhart and Russell, 1966))

However, Peterson *et al.* (1992) and Fasahat *et al.* (2015) reported that the concept of optimal genotype stability and response to environments for quality parameters differs somewhat from that conventionally used to describe yield stability. For breeders, stability of quality attributes is important from the points of changing genotypes ranks' throughout environments and influences selection efficiency. For end users, stability in quality properties of genotypes is more important, irrespective of genotypes rank changes. However, as pointed out by Grausgruber *et al.*, (2000) the quality of a genotype often behaves similar to other quantitative characters to desirable and undesirable environmental conditions. As a result, a genotype is regarded stable if it has a low contribution to the GE interaction.

Stability of desirable genetic characters is important for development of improved varieties and useful for the commercial exploitation over a wide range of agro-climatic conditions. A commercial variety must have stable performance and broad adaptation over a range of environments in addition

to high yield potential. It is more practical to develop and release varieties which are adapted to more than a single environment and can be successfully grown over a range of environments. For this, a variety well adapted to more than one environment and stable has to be selected. Preliminary evaluation can be done to identify stable genotypes through screening (Nahar *et al.*, 2010).

Fikere *et al.*, (2008) and Nahar *et al.*, (2010) stated that in addition to high yield potential, a new cultivar should have stable performance and broad adaptation over a wide range of environments. Evaluating stability of performance and range of adaptation has become increasingly important for breeding programs. Hence, if cultivars are being selected for a large group of environments, stability and mean performance across all environments are important than performance for specific environments (Piepho, 1996).

2.12 Stability Analysis

Stability analysis provides a general summary of the response patterns of genotypes to environmental change (Albert, 2004)

The success of any breeding program depends on several factors, including understanding and selection of suitable breeding test locations (Dia, 2012). Measuring G x E interaction is very important in determining an optimal breeding strategy for releasing genotypes with an adequate adaptation to target environments (Fox *et al.*, 1997).

Breeders usually look for a variety that has good mean trait performance over a wide array of environments and years and the concept of stability is overlooked. Such an approach is reasonable if there are no genotype by environment interactions, but in most cases interactions do exist. Some genotypes can have high yield in few environments and very low yield in other environments, showing better mean performance across environments. But few genotypes may have average performance that is stable over wider environments (Magagane, 2012). Hayward *et al.*, 1993 stated that knowledge of the pattern and magnitude of genotype by environment interactions and stability analysis is important for understanding the response of different genotypes to varying environments and for identifying superior soybean genotypes under the target environment and agronomic conditions to maximize specific adaptation and to speed up the transfer of new cultivars to growers.

The advantage of selecting superior genotypes using stability analysis instead of average performance is that stable genotypes are dependable across the environments which reduces G x E

Interaction. Studies have shown that stability analysis according to various principles can result in better identification of stable genotypes, even when there were no interactions among the parameters (Fasahat *et al.*, 2015).

Stability analysis provides a general summary of the response patterns of genotypes to environmental change. Various statistics have been proposed to measure the stability of genotypes over environments (Dia, 2012). The main problem with stability statistics is that they don't provide an accurate picture of the complete response pattern (Hohls, 1995). This is because a genotype's response to varying environments is multivariate (Lin *et al.*, 1986) whereas the stability indices are usually univariate (Gauch, 1988; Crossa, 1990; Odewale *et al.*, 2013).

Freeman (1973) termed the main type of stability analysis as joint regression analysis or joint linear regression (JLR). It involves the regression of the genotypic means on an environmental index. Joint regression analysis provides a means of testing whether the genotypes have characteristic linear responses to changes in environments. Joint regression analysis was first proposed by Yates and Cochran (1938) and then widely used and reviewed by various authors (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966; Perkins and Jinks, 1968; Wright, 1971; Freeman and Perkins, 1971; Shukla, 1972; Hardwick and Wood, 1972; Freeman, 1973; Hill, 1975; Lin *et al.*, 1986; Westcott, 1986, Becker and Léon, 1988; Baker, 1988; Crossa, 1990; Hohls, 1995).

Other statistical methods for stability analysis include; Multivariate ANOVA, multiple regression, Principal Component Analysis (PCA), Factor analysis, clustering and ordination and Additive main Effect and Multiplicative Interaction (AMMI) model. Since the genotype response to environmental variations is usually multivariate, a multivariate method of analysing genotype stability across environments is the best option (Odewale *et al*, 2013). Multivariate statistical methods explore multidirectional parameters and extract more information on the components of phenotypic variability (Hussein, 2000; Nkhoma, 2013). One of the multivariate statistical techniques is the Additive Main Effects and Multiplicative Interaction (AMMI) model.

2.13 AMMI

The Additive Main Effects and Multiplicative Interaction (AMMI) model was first introduced in social sciences and physics (Mandel, 1961, 1969; Gollob, 1968; Crossa, 1990), and was later adapted to the agricultural context (Gauch, 1988, 1992; Cornelius, 1993; Piepho, 1996). This model was considered appropriate if one is interested in predicting genotypic yields in specific environments

(Annicchiarico, 1997). It combines the analysis for the genotype and environment main effect with several graphically represented interactions for principal component analysis (IPCAs) (Crossa, 1990; Abamu and Alluri 1998). Thus, it helps to summarizing the pattern and relationship of genotypes, environment and their interaction (Gauch and Zobel, 1996).

The AMMI model combines both classical ANOVA and PCA into a single model with additive and multiplicative parameters (Zobel *et al.*, 1988; Shafii and Price, 1998; Pinnschmidt and Hovmoller, 2002). The model separates the additive variance from the interaction variance and applies PCA to the interaction portion from the ANOVA analysis to extract a new set of coordinate axes that account more effectively for the interaction patterns (Zobel *et al.*, 1988; Shafii and Price, 1998; Thillainathan and Fernandez, 2001). In clarification of GE Interaction, AMMI summarises patterns and relationships of genotypes and environments (Crossa, 1990). Furthermore, statistical model results from AMMI analysis are plotted in a graph showing the main and interaction effects for both genotypes and environments on the same scatter plot, with the noise rich residual discarded and the data separated into a pattern rich model to gain accuracy (Gauch and Zobel, 1996)

Among the various statistical procedures developed for the study of G x E interaction, the AMMI model has been revealed to be efficient because it captures a large portion of the GE sum of squares and austerely separates main and interaction effects that present agricultural researchers with different kinds of opportunities, and this model often provides agronomically meaningful interpretation of the data (Ebdon and Gauch, 2002). The results of AMMI analysis are useful in supporting breeding program decisions such as specific adaptations to target environments (tolerances to disease, heat and drought, cold) and selection of environments or test site locations (Gauch and Zobel, 1997; Riaz *et al.* 2013).

The AMMI method is being used in studies on the G x E interaction and stability of soybean (Zobel *et al.* 1988; Oliveira *et al.* 2003).

2.13.1 AMMI Biplots

The AMMI biplots are used to visualise the adaptability (average performance across localities) and stability (consistency) of genotypes. A graphically represented AMMI analysis enables selection of stable and high yielding cultivars for a given region, as well as cultivars with specific adaptability. The differences and genotype distributions in the biplot are a consequence of genotype variations in different conditions (Marjanovic – Jeromela *et al.* 2011).

The IPCA scores of a genotype in the AMMI analysis are an indication of the stability or adaptability of a genotype over environments (Gauch and Zobel, 1996). The greater the IPCA scores, negative or positive (as it is a relative value), the more specifically adapted a genotype is to certain environments. The more the IPCA scores approximate zero (0), the more stable the genotype is over all the environments sampled. If the IPCA scores of a genotype are interpreted in conjunction with the IPCA scores of the individual environments (AMMI 1 biplot), the adaptability of the genotype can largely be determined by characterization of the environments, for example whether they are low potential environments (Schoeman, 2003). The AMMI 1 biplot shows both main and interaction effects for both genotypes and environments. The abscissa shows the main effects (means) while the ordinate shows the effects of the first interaction principle axes (IPCA1) scores.

When there is interest in more than just the first IPCA axis, the AMMI 2 biplot is used with IPCA 1 on the abscissa and IPCA 2 on the ordinate. Unlike AMMI 1 biplot which does not show the additive main effects, by showing two IPCA axes, the AMMI 2 biplot captures a higher percentage of the interaction (for IPCA1 & IPCA2) compared with only one for IPCA 1. When IPCA 2 is sizable and significant, such a graph is a useful supplement to the usual biplot and can be used to further explore adaptation.

The AMMI 2 biplot explains the magnitude of the interaction of each genotype and environment. The genoypes and environments that are furthest from the origin being more responsive fit the worst. Genotypes and environments that fall into the same sector interact positively while those that fall into opposite sectors interact negatively (Osiru *et al.*, 2009). A genotype showing high positive interaction with an environment obviously has the ability to exploit that agro-ecological or agromanagement conditions of the specific environment and is therefore best suited to that environment. AMMI analysis permits estimation of interaction effect of a genotype in each environment and it helps to identify genotypes best suited for specific environmental conditions. (Rashidi *et al.*, 2013).

2.13.2 AMMI Stability Value

The AMMI model does not provide for a quantitative stability measure but such a measure is essential to quantify and rank genotypes in terms of stability (Gauch, 1992; Gauch and Zobel, 1996). Therefore Purchase *et al.* (2000), proposed the AMMI stability value (ASV) to quantify and rank genotypes according to their stability. The ASV is the distance from zero in a two- dimensional scatter gram of IPCA1 scores and IPCA2 scores. Since the IPCA1 score contributes more to GE sums of squares, it has to be weighted by the proportional difference between IPCA 1 and IPCA 2

scores to compensate for the relative contribution of IPCA 1 and IPCA 2 to the total GE sum of squares. The distance from zero is then determined using Pythagorean Theorem (Purchase *et al.*, 2000). In the ASV method, the larger the ASV value, the more specifically adapted a genotype is to certain environments. Smaller ASV scores indicate a more stable genotype across environments.

$$ASV = \sqrt{\left[\frac{SS \text{ IPCA 1}}{SS \text{ IPCA score 1}}\right]^2 + [IPCA \text{ score}]^2}$$

Where:

SS = Sum of squares;

 $(SS_{IPCA1}) \div SS_{IPCA2}) =$ the weight given to the IPCA1-value by dividing the IPCA1 sum of squares by the IPCA2 sum of squares.

2.14 Near Infrared Reflectance Spectroscopy

Near Infrared Reflectance (NIR), discovered by Friedrich Wilhelm Herschel in 1800 is a rapid, nondestructive, inexpensive and accurate method for simple and rapid analyses of various agricultural and food products. NIR covers the range of electromagnetic spectrum from 780 to 2500nm. In NIR spectroscopy, the product to be tested is irradiated with NIR, and the reflected or transmitted radiation is measured. The spectral characteristics of the material being tested changes as the radiation penetrates it by means of wavelength dependent scattering and absorption processes. The change is dependent on the chemical composition of the product being tested and on its light scattering properties which are related to the microstructure (Lee *et al.*, 2011).

Over the years, NIR has been widely used for analysis of soybean constituents (Sato *et al.*, 2008). In 1974, Hymowitz used NIR for determination of oil and protein of soybean seed. Choung et al., 2001, Sato *et al.*, 2011 and Baianu *et al.*, 2010 also employed NIR for determination of soybean seed oil and protein content. Use of NIR in measuring characteristics can greatly enhance progress in improving soybean for important seed components (Lee *et al.*, 2011).

2.15 Plant Protein Composition

Proteins are large molecules that consist of long chains of amino acids covalently joined through peptide bonds. Proteins serve numerous functions both within and outside cell and these include

structural roles, catalysing chemical reactions (enzymes), and facilitating membrane transport and energy generating reactions involving electron transport.

There are 20 different known amino acids and each of these amino acids has a fundamental design consisting of a central carbon bonded to a hydrogen (-H), a carboxyl group (-COOH), an amino group (-NH₂) and a unique side chain or R- group (Suri, 2006; Gallagher, 2010). The distinguishing characteristic between amino acids is the unique side chain or R- group as it is the chain that dictates an amino acid's chemical properties.

2.16 Some Important Elements in Amino Acid and Protein Formation

2.16.1 Sulphur

Sulphur is a constituent of the sulphur containing amino acids namely cysteine and methionine as well as other metabolites. Sulphur along with nitrogen is required in the synthesis of these proteins and some enzymes (Jamal *et al.*, 2005). Sulphur also plays a vital role in various plant growth and development processes.

2.16.2 Iron

Iron (Fe) is required in legume plants for important processes such as energy transfer, respiration, photosynthesis, DNA synthesis, and nitrogen fixation. Iron is important for nodulation which is required for accumulation of nitrogen in plants. Nitrogen is in turn a constituent of many compounds including amino acids, proteins and nucleic acids (Marschner, 2011) and its uptake may influence the concentration of protein in grains. Legumes with an active symbiosis have a large requirement for iron because several symbiotic proteins incorporate iron. Numerous bacteroids require iron for synthesis of nitrogenase, the nitrogen-fixing enzyme, as well as cytochromes, hydrogenase and ferredoxin (O'Hara, 2001; Peters and Szilagyi, 2006). According to Hemantarajan and Trivedi (1997) application of iron increases protein content in soybean seed. N deficiency can be one of the side effects resulting from iron deficiency as iron is essential for nodule formation and function. Iron deficiency can result in reduced rates of nitrogen fixation and could affect nodule initiation and development.

2.16.3 Phosphorus

Phosphorus (P) is the second major nutrient element that is essential for crop growth. The most obvious effect of P is on the plant root system. Since P plays a very important role in the formation
of nodules and nitrogen fixation, nodulating legumes require higher amounts of P compared to nonnodulating crops (Kabir *et al.*, 2013). P deficiency therefore impairs nodulation as it affects the assembly of functional iron- sulphur cluster (Burton, *et al.*, 1998).

P is also important in the development of an extensive root system (Sharma and Yadav, 1997, Gobarah et al., 2006) and therefore enables plants to absorb more water and nutrients from depth of the soil and this in turn enhances the plant's ability to produce more assimilates.

The role of P in building phospholipids and nucleic acid is known as it is involved in activation of metabolic processes. P is a very important nutrient for all crops and particularly for legumes, as it is a major constituent of ATP and plays an important role in energy transformation in plant and seed formation (Kabir *et al.*, 2013)

CHAPTER 3 MATERIALS AND METHODS

3.1 Trial Locations

The experiments were conducted during the 2013/ 2014 rain season at five selected locations of Zambia's three agro- ecological zones namely Masumba Research Sub- Station in Mambwe, Msekera Research Station in Chipata, Kabwe Research Station in Kabwe, Golden Valley Agriculture Research Trust (GART) in Chisamba and Misamfu Research Station in Kasama (Figure 1). A general description of the locations used is provided in Table 1a.

Zambia is sub-divided into three agro- ecological regions based on climatic characteristics with rainfall being the main factor (Bunyolo *et al.*, 1997). Masumba falls under Agro- ecological Region I of Zambia, Msekera, Kabwe and GART fall under Region II, while Misamfu falls under Region III.

Region I mainly covers the valleys lying between 300 and 900 metres above sea level (Low altitude region), and receives mean annual rainfall of not more than 800mm ranging from 80 to 120 days at 70% probability. Region I experiences up to five 10 day dry periods of less than 30mm of rainfall and is the driest and most prone to drought. Relatively high temperatures characterize this region with mean daily temperatures varying from 20 to 25°C during the growing season. Region II (Mid-altitude region) with elevations between 900 to 1300 metres above sea level receives medium mean annual rainfall of between 800 and 1000 mm. The rainfall is generally well distributed at 70% probability and ranges from 100 to 140 days and may contain one to three 10 day dry periods of less than 30 mm rainfall. Mean daily temperatures during the growing season range from 23°C to 25°C. Region III (High altitude region) with altitudes ranging from 1100 to 1700 metres above sea level receives mean annual rainfall exceeding 1000 mm with the growing season ranging from 120 to 150 days and rarely experiences any drought. Mean monthly temperatures during the growing the growing season in region III range from 16°C to 24°C (Bunyolo *et al.*, 1997).

Average climatic conditions at the locations are presented in Table 1a. The wettest location was Misamfu while Kabwe was the driest. Temperatures were coolest at Misamfu and hottest at Masumba. GART and Kabwe were similar in amount of average rainfall received and the mean temperature, however the two locations had very different soil type. Msekera had the second highest average rainfall but also experienced the second coolest temperatures. The 2013/2014 growing

season as shown in Table 1b reveals that the season was generally normal and experienced favorable temperatures and good rains which were well distributed throughout the growing season apart from Kabwe and GART which experienced low rainfall towards the end of the season. Soil analysis results for the five trial sites are presented in Table 1c.



Figure 1: Map of Zambia showing the five experimental sites

Source: Google Maps

Agro- Ecological Region	Site Name	Lat. ° S	Long. ° E	Alt. (m asl)	Temp °C	рН	Mean Annual Rainfall (mm)	Soil texture
Ι	Masumba	13.22143	31.92812	791	32.88	5.52	642.8	Sandy loam
Π	GART, Chisamba	14.49684	28.09979	1148	24.24	5.95	601.5	Sandy loam
II	Kabwe	14.39472	28.49474	1177	23.12	5.52	583.3	Loamy sand
II	Msekera	13.64658	32.56085	1104	29.5	5.56	1,097.7	Sandy loam
III	Misamfu	10.17307	31.22203	1400	21.66	4.62	1,348.4	Loamy sand

Table 1a: General Features of the Trial Sites with soil and climatic data during the 2013/2014 season

Table 1b: Monthly Meteorological data	of sites used in the study	during the 2013/2014 season
---------------------------------------	----------------------------	-----------------------------

		Month							
Location		Dec	Jan	Feb	Mar	Apr			
Masumba	Rainfall (mm)	106.9	246.3	214.1	75.5	0			
	Mean Temp (°C)	35.6	31.8	31.8	33	32.2			
Msekera	Rainfall (mm)	143.1	306.5	407.8	216.8	23.5			
	Mean Temp (°C)	31.6	28.5	28.5	30.1	28.8			
Kabwe	Rainfall (mm)	191.7	204.2	97	88.4	2			
	Mean Temp (°C)	24.9	23.5	22.95	22.95	21.3			
GART	Rainfall (mm)	307.6	69.2	99.4	65.1	60.2			
	Mean Temp (°C)	25.2	25.1	24.4	24.1	22.4			
Misamfu	Rainfall (mm)	315.9	234.4	464.1	256.3	77.7			
	Mean Temp (°C)	21.9	21.75	21.5	21.95	21.2			

Sample	рН	Ν	Org. Matter	Р	K	Na	Ca	Mg	Cu	Fe	Mn	Zn	S
	0.01M		Walkley & Black	Bray 1	Amm Ac	etate			DTPA	A			Na Acetate
	CaCl ₂	CaCl ₃	%	mg/kg	cmol/kg				mg/kg	5			
Kabwe	5.52	0.06	0.56	15.21	0.17	0.05	1.83	0.57	0.14	6.44	6.43	0.6	14.79
GART	5.95	0.07	1.92	7.56	0.66	0.08	6.5	2.47	3.24	3.38	6.26	0.9	17.75
Misamfu	4.62	0.22	1.68	11.58	0.16	0.06	0.82	0.36	0.05	10.2	3.34	0.3	23.18
Msekera	5.63	0.08	2.4	12.27	0.9	0.1	10.00	2.25	0.64	9.46	8.03	0.7	13.81
Masumba	5.52	0.07	3.52	1.99	0.43	0.06	6.83	1.51	0.97	6.92	9.61	0.6	12.82

Table 1c: Soil Analysis Results for sites used in the study

3.2 Genotypes used in the study

Twelve (12) soybean genotypes representing a fixed sample were selected from the International Institute of Tropical Agriculture (IITA)'s ongoing soybean improvement program, two (2) improved varieties from the Zambia Agriculture Research Institute (ZARI) and one (1) from SeedCo Zambia Limited were used in the study (Table 2). These materials obtained from Zambian institutions served as checks. The pedigree information of the IITA genotypes is presented in Figure 2 however the pedigree for TGx 1904- 6F could not be verified and therefore was not included.

Code	Genotype	Characteristic	Source
G1	TGx 1740-2F	Promiscuous	IITA
G2	TGx 1830-20E	Promiscuous	IITA
G3	TGx 1835-10E	Promiscuous	IITA
G4	TGx 1887-65F	Promiscuous	IITA
G5	TGx 1904-6F	Promiscuous	IITA
G6	TGx 1987-11F	Promiscuous	IITA
G7	TGx 1987-23F	Promiscuous	IITA
G8	TGx 1988-9F	Promiscuous	IITA
G9	TGx 1988-18F	Promiscuous	IITA
G10	TGx 1988-22F	Promiscuous	IITA
G11	TGx 1989-60F	Promiscuous	IITA
G12	TGx 1990-129F	Promiscuous	IITA
G13	Magoye (check)	Promiscuous	ZARI
G14	Safari (check)	Non Promiscuous	SEEDCO
G15	Lukanga (check)	Non Promiscuous	ZARI

Table 2: Genotypes used in the Study



Source: IITA, 2013 Figure 2: Pedigree of IITA genotypes used in the study

3.3 Experimental Design and Management

3.3.1 Field Layout and Design

The genotypes were arranged in a Randomized Complete Block Design (RCBD) with four replications, at all the five trial sites. Gross plot size was 4 rows, 50 cm apart and 6m long while the net plot size was 2 rows, 50 cm apart and 6m long.

3.3.2 Trial Planting and Cultural Practices

The trials were planted between 18^{th} and 24^{th} December, 2013 as shown in Table 3. The planting was done by hand on flat land at all the sites at a seed rate of 80kg/ha (96g per plot). Basal fertilizer applied by broadcasting at planting was Compound D (10% NO₄; 20% P₂O₅ and 10% K₂O) at a rate of 100kg/ha.

Weeds were managed through hand weeding. Scouting for diseases as well as insect pest attack was also done regularly at various growth stages. Fortnight spraying with Cypermethrin and Acetamiprid was used to control insects while difenoconazole fungicide was used for rust control.

The crop was rain fed with no supplementary irrigation. Hand harvesting of the crop was done when 95% of the leaves had turned yellow to brown and were dropping off.

	Location	Code	Planting Date
1	Kabwe	E1	18 th December, 2013
2	Misamfu	E5	20 th December, 2013
3	Msekera	E3	23 rd December, 2013
4	Masumba	E4	24 th December, 2013
5	GART	E2	27 th December, 2013

Table 3: Planting Dates for the five Locations

3.4 Data Collection

As this study focused on seed variables, protein and oil, seed was collected at harvest time (physiological maturity) from each net plot. The soybean crop reaches physiological maturity when 95% of pods reach their mature pod color which can be brown, gold, yellow or grey depending on the variety. The beans in the pods have shrinked and separated from the white membrane inside the pod at physiological maturity (Kandel, 2015). The field trial had four replications, however for the

current study, only two replications were used for analysis due to the inhibiting cost of determining the oil and protein content. Samples of 35 g were collected and analyzed for chemical composition. Protein and oil concentration was determined by the Department of Crop Sciences at University of Illinois using a Perten DA7200 Diode Array Near Infrared Reflectance (NIR) analyzer with built in calibration. Laboratory analysis of soil samples was done by the UNZA School of Agriculture soil science.

3.5 Statistical Analysis

GENSTAT Statistical package version 14 was used for the analysis of variance (ANOVA) for each of the measured and derived parameters as well as for the Additive Main Effect Multiplicative Interaction (AMMI) Model (Gauch, 1992). Simple correlation analysis was used to establish the association/relationship between soybean oil content and the various climatic and soil parameters as well as protein content and climatic and soil parameters. Minitab statistical package version 14 was used for stepwise regression.

3.5.1 Analysis of variance

The oil and protein content data were subjected to analysis of variance (ANOVA) separately for each location to establish if there were any statistical differences in the performance of these genotypes for oil and protein content. A mixed model was adopted. Furthermore, a combined across locations ANOVA was done in order to determine differences between genotypes across locations and also to determine whether there was significant difference among the locations. Mean separation was done using Least Significant Difference (LSD) procedure at 5% probability level (Gomez and Gomez, 1984; Steele *et al.*, 1997) to discriminate the genotypes and identify superior ones based on the trait of interest.

The statistical model for ANOVA of a randomized complete block design used was:

 $Yij = \mu + \beta j + Ti + \varepsilon ij,$

Where;

 μ = the overall mean,

 βj = the jth block effect,

Ti = the ith treatment effect, and

Cij = the experimental error.

The statistical model for combined ANOVA used was;

 $Yijk = \mu + Gi + Ej + GEij + Bk (j) + eijk$

Where;

Yijk = observed value of genotype i in block k of environment (location) j,

 μ = the overall mean,

Gi = effect of genotype i,

Ej= environment or location effect,

GEij = the interaction effect of genotype i with environment j,

Bk (j) = the effect of block k in location (environment) j,

eijk= error (residual) effect of genotype i in block k of environment j.

3.5.2 Correlation Analysis

Simple correlation analysis was used to establish the associations among soybean oil content and the various climatic and soil parameters as well as protein content and climatic and soil parameters.

3.5.3 Multiple Linear Regression

A stepwise multiple regression (backward and forward) was conducted in order to study the cause and effect relationship between the measured variables. In this analysis, oil and protein were the dependent variables while climatic and edaphic characteristics measured in the study were the independent variable.

Regression analysis establishes the cause-effect relationship as it helps to explain the variation in the dependent variable using the independent variables. In this study, regression analysis was used to establish the key climatic and/or edaphic factors that were important in explaining the variation in

the dependent variables. Using results from this analysis, it was possible to then explain the changes in the dependent variables due to the independent ones. This analysis also allowed for insight into the nature of what was generally defined as environment

3.5.2 Stability Analysis

After the test of significance for GEI in the ANOVA, a stability analysis was conducted for oil and protein content using AMMI. The AMMI model combines the ANOVA for the genotype and environment main effects with PCA of the G x E interaction. The scores or coordinates of the genotypes and the environments are produced on the principal interaction axes conventionally called IPCA, that permit their representation together in a biplot graph. In the AMMI analysis, locations and replications were considered random effects, whereas genotypes were considered as fixed effects.

The AMMI model equation used was;

$$Y_{ger} = \mu + \alpha_g + \beta_e + \Sigma_n \lambda_n \gamma_{gn} \delta_{en} + \rho_{ge} + \epsilon_{ger}$$

Where;

 Y_{ger} = the yield of genotype g in the environment e for replication r,

 μ = the grand mean

 α_g = the deviation of the genotype g from the grand mean,

 β_e = the deviation of the environment e from the grand mean

 λ_n = the singular value for the interaction principal component axis (IPCA) n,

 γ_{gn} = the genotype eigenvector value for IPCA axis n, (square root of the eigen value which is also the sum of squares divided by the number of replications),

 δ_{en} = the environment e eigenvector vector value for IPCA axis n,

 ρ_{ge} = the residual and

 ϵ_{ger} = the error term if the experiment is replicated.

The eigen vectors scaled as unit vectors are unit less, while μ , αg , βe are additive parameters and enter the model additively and $\lambda_n \gamma_{gn}$ and δ_{en} are multiplicative parameters that enter the model multiplicatively (Gauch and Zobel, 1996).

To investigate the main effects and interactions, AMMI1 Biplot was constructed for protein content for the fifteen genotypes across five environments. The IPCA 1 scores were plotted against IPCA 2 scores in the AMMI 2 biplot to further explore adaptation.

CHAPTER 4 RESULTS

4.1 Soybean oil and protein content

4.1.1 Oil Content

The analyses of variance for oil content on a location basis revealed significant differences ($p \le 0.05$) among genotypes at all locations except Misamfu (Table 4). The mean performance of the genotypes varied from one location to another (Table 5). At Kabwe the range of performance was from 19.6% being highest (Lukanga) to 15.54% being the lowest (genotype TGx 1987-23F). Among the IITA genotypes the highest oil content was that of TGx 1988-22F (19.24%) with genotype TGx 1987-23F having the lowest oil content (15.54%). Average oil content of the checks was 19.09% compared to the IITA genotypes' average of 17.77%. The best check was Lukanga (19.6%).

Oil content at GART on the other hand ranged from 19.3% (Lukanga-check) being the highest to 15.05% (TGx 1987-23F) being the lowest. TGx 1988-22F had the highest oil content among the IITA genotypes (17.12%) while TGx 1987-23F had the lowest oil content (15.05%) among the IITA genotypes. The checks had an average oil content of 17.85% at GART compared to the IITA genotypes which had a lower average of 16.02%. The best check was Lukanga (19.3%) and this was also the highest oil yielding genotype among all the entries at this site.

At Msekera, the oil content ranged from 21.09% to 17.04% with Lukanga (check) and TGx 1830-20E being the highest and lowest oil yielders respectively. Among the IITA genotypes, TGx 1988-22F had the highest oil content (19.87%) while TGx 1830-20E had the lowest oil content (17.04%). The checks had a higher average oil content of 20.14% than the IITA genotypes which had an average oil content of 18.69%. The best check at this site was Lukanga (21.09%) and was also the highest yielding genotype with respect to oil content amongst all tested genotypes at this site.

The mean performance for oil content at Masumba ranged from 19.58% (TGx 1988-9F) to 16.61% (TGx 1830-20E). TGx 1988-9F had the highest oil content (19.58%) among the IITA genotypes while TGx 1930-20E (16.61%) had the lowest oil content among the IITA genotypes at Masumba. The checks had an average oil content of 18.57% compared to the IITA genotypes which had an average oil content of 17.98%. The best check at this site was Lukanga (19.3%).

		Kabw	e	(GART	l	Msekera	M	asumba		Mis	samfu
Source	d.f.	(m.v.)	m.s.	d.f.	m.s.	d.f.	m.s.	<u>d.f.</u>	m.s.	<u>d.f.</u>	(m.v.)	m.s.
Rep	1		0.522	1	0.6901	1	0.07008	1	0.9612	1		2.0991
Genotype	14		2.5266* *	14	2.8067**	14	2.20372**	14	1.3496*	14		0.5749 ^{NS}
Error	12	(2)	0.1265	14	0.256	14	0.07953	14	0.2999	11	(3)	0.8696
Total	27	(2)		29		29		29		26	(3)	
Overall Mea	n		18.03		16.38		18.98		18.13	36.9		18.86
CV			2		3.1		1.5		3	3.1		4.9

 Table 4. Analysis of variance per location for soybean oil content of 15 soybean genotypes

 NS = Non significant, * = significant at p \leq 0.05, ** =Significant at p \leq 0.001 respectively, CV= Coefficient of variation, d.f. = degree of freedom, m.s. = Mean Square

				Oil %		
Genotype	Kabwe	GART	Msekera	Masumba	Misamfu	Overall Genotype Means
TGx 1740-2F	18.44cde	17.03abc	19.31cde	17.79ab	19.19a	18.35cde
TGx 1830-20E	15.75ab	15.28a	17.04a	16.61a	18.96a	16.73a
TGx 1835-10E	17.59bcd	15.17a	17.8ab	17.21ab	18.48a	17.25ab
TGx 1887-65F	17.38bc	16.09ab	18.71bcde	17.43ab	19.09a	17.74bcd
TGx 1904-6F	18.23cde	16.21ab	19.63de	18.18ab	18.18a	18.09cde
TGx 1987-11F	18.27cde	16.25ab	19.68de	18.13ab	19.64a	18.39cde
TGx 1987-23F	15.54a	15.05a	17.65ab	17.49ab	19.7a	17.09ab
TGx 1988-9F	18.2cde	15.2a	18.46bcd	19.58b	19.21a	18.13cde
TGx 1988-18F	18.21cde	16.58ab	18.8bcde	18.73ab	18.36a	18.13cde
TGx 1988-22F	19.24de	17.12abc	19.87ef	18.75ab	18.19a	18.63def
TGx 1989-60F	18.21cde	16.57ab	19.27cde	18.29ab	19.01a	18.27cde
TGx 1990-129F	18.12cde	15.67a	18.07abc	17.59ab	18.67a	17.62abc
Magoye (check)	18.65cde	16ab	19.53de	17.97ab	19.5a	18.33cde
Safari (check)	19.03cde	18.24bc	19.79e	18.98ab	18.71a	18.95ef
Lukanga (check)	19.6e	19.3c	21.09f	19.3b	18.09a	19.47f
Location Means	18.03b	16.38a	18.98c	18.13b	18.86c	18.08
Min	15.54	15.05	17.04	16.61	18.09	16.73
Max	19.6	19.3	21.09	19.58	19.7	19.47
Mean of IITA entries	17.77	16.02	18.69	17.98	18.89	17.87
Mean of checks	19.09	17.85	20.14	18.75	18.77	18.92
LSD 5%	0.5481	0.7673	0.4277	0.8306	1.451	0.495
CV	2	3.1	1.5	3	4.9	

 Table 5: Mean Oil content (%) of fifteen soybean genotypes tested at the five different locations during the 2013/2014 season

		Oil Content (%)							
Source of variation	d.f.	(m.v.)	m.s.						
Location	4		32.3452**						
Reps (L)	5		0.8682						
Genotype	14		5.0774**						
Location*Genotype	56		1.096**						
Residual	65	(5)	0.3074						
Total	144	(5)							

 Table 6: Combined analysis of variance across five locations for soybean oil content.

*, ** Significantly different at $p \le 0.05$ and $p \le 0.001$ levels respectively, d.f. = degree of freedom, m.s. = Mean Square

Combined analysis of variance for oil content showed significant differences ($p \le 0.01$) among Locations and among Genotypes (Table 6). Significant Genotype by Location interactions for oil content were also observed.

The locations were highly significantly different from each other for oil content across genotypes with Msekera having the highest location mean (18.98%) among the five locations and GART having the lowest location mean (16.38%) (Table 5). Similarly, the genotypes were significantly different for oil content when all locations were considered with Lukanga (a check), having the highest across location average oil content (19.47%) and TGx 1830-20E as the lowest oil performing genotype with 16.73% across locations.

The combined analysis of variance also revealed that all the three checks used in the study had oil content above the overall mean for all the genotypes (18.08 %) with Safari and Magoye ranking second and sixth respectively among all the genotypes tested at all five locations.

Analysis of variance for oil content also revealed significant GE interactions (Table 6). From table 5, it was observed that genotypes changed in both oil content magnitudes and ranking. The nonconsistent performance of the genotypes manifested interactions with environments. For instance, the best genotype at Masumba (TGx 1988-9F) was fourth at Misamfu, tenth at Kabwe, thirteenth at GART and eleventh at Msekera. Similar fluctuations were observed for TGx 1988-22F which was the second best performing genotype at Kabwe and Msekera; its ranking fluctuated from being third and fourth at GART and Masumba, respectively and was thirteenth at Misamfu. The check Lukanga was the best performer at Kabwe, GART and Msekera and was second best at Masumba. Its ranking fluctuated to fifteenth at Misamfu where it was the lowest performer. The check Magoye also showed fluctuations with its ranking fluctuating from fourth at Kabwe to tenth, sixth, ninth and third at GART, Msekera, Masumba and Misamfu respectively.

Simple correlation analysis of oil content with climatic and edaphic parameters (Appendix 1) showed simple relationships between oil and iron (r = 0.84), oil and rainfall (r = 0.50), oil and calcium (r = -0.84) and oil and Zinc (r = -0.50)

A stepwise regression analysis to adduce the main causal factors to the fluctuations of oil content in soybean revealed that two factors namely, iron (Fe) and rainfall, being edaphic and climatic factors respectively, were the most important causal factors with b-values of 0.53 and -0.00145 respectively and explaining up to 48% of the variation in oil content ($R^2 = 47.78$). The sensitivity of oil to change in rainfall is implied by the magnitude of the associated b- value (b = - 0.00145). These results also showed that the other factors were not important determinants, in nature, to the changes in oil content.

The resulting regression equation for oil was;

Oil = 15.4 + 0.531 Fe - 0.00145 Rainfall R2 adj = 47.78%

4.1.2 Protein Content

The analyses of variance per individual location for protein content revealed significant differences $(p \le 0.05)$ among genotypes for protein content at all locations but Misamfu (Table 7).

	Kabwe		bwe	GART		Msekera		Masumba		Misamfu		
Source	d.f.	(m.v.)	m.s.	d.f.	m.s.	d.f.	m.s.	d.f.	m.s.	d.f.	(m.v.)	m.s.
Rep	1		0.9884	1	1.5916	1	0.0154	1	0.031	1		7.708
Genotype	14		7.0012**	14	4.68885**	14	8.1378**	14	4.209*	14		5.816 ^{NS}
Error	12	(2)	0.4648	14	0.08323	14	0.8028	14	1.341	11	(3)	9.128
Total	27	(2)		29		29		29		26	(3)	
Overall M	ean		36.34		38.23		34.55			36.9		33.47
CV%			1.9		0.8		2.6			3.1		9

Table 7. Analysis of variance per location for soybean protein content of 15 soybean genotypes

 $\overline{^{NS}}$ = Non significant, * = significant at p \leq 0.05, ** =Significant at p \leq 0.001 respectively, CV= Coefficient of variation, d.f. = degree of freedom, m.s. = Mean Square

At Kabwe (Table 8) the range of performance was from 39.29% being highest (TGx 1987-23F) to 32.56% being the lowest (Lukanga). TGx 1987-23F also had the highest protein content (39.29%) among the IITA genotypes while TGx 1740-2F had the lowest (35.36%). The average protein Content of the checks was 33.7% compared to the IITA genotypes' average of 39.29%. The best check at Kabwe was Safari (34.66%) while the lowest was Lukanga (32.56%).

Protein content at GART on the other hand ranged from 40.12% (TGx 1988-9F) to 33.55% (Lukanga) being the highest and lowest respectively. Among the IITA genotypes the highest protein content was of TGx 1988-9F (40.12%) with genotype TGx 1988-18F having the lowest (37.68%). At this site, the average protein content of the checks was 36.03% compared to the IITA genotypes' average of (38.77%). The best check was Magoye (37.66%) while the lowest check was Lukanga (33.55%).

At Msekera, the protein content ranged from 37.02% being the highest and 29.61% being the lowest for the genotypes TGx 1830-20E and Lukanga (check) respectively. TGx 1830-20E (37.02%) had the highest protein content among the IITA genotypes while the lowest among the IITA genotypes was TGx 1904-6F (32.44%). The checks had an average protein content of 32. 73% and the best performing check was safari (34.6%) while the lowest was Lukanga (29.61%).

At Masumba, the range of performance for protein content was from 39.16% (TGx 1988-22F) to 34.1% (Lukanga). Among the IITA genotypes, TGx 1988-22F had the highest protein content (39.16%) while TGx 1988-9F was the lowest (35.43%). The average protein content for the checks was 35.13% with Magoye being the best check (36.06%) and Lukanga being the lowest (34.1%).

			Protein %	0		
Genotype	Kabwe	GART	Msekera	Masumba	Misamfu	Overall Genotype Means
TGx 1740-2F	35.36abcd	38.23bcd	34.08bcd	37.19ab	33.11a	35.59bcd
TGx 1830-20E	38.87ef	39.02cd	37.03d	39.15b	33.8a	37.57d
TGx 1835-10E	35.39abcd	37.77bc	33.52bcd	36.27ab	29.12a	34.41abc
TGx 1887-65F	37.86cdef	39.03cd	36.05bcd	38.1ab	35.39a	37.28cd
TGx 1904-6F	35.9bcde	38.62bcd	32.44ab	36.05ab	34.17a	35.44abcd
TGx 1987-11F	36.78bcdef	39.34cd	32.67abc	38.25ab	31.95a	35.8bcd
TGx 1987-23F	39.29f	39.23cd	36.56d	37.9ab	31.91a	36.98bcd
TGx 1988-9F	38.31def	40.12d	36.34cd	35.43ab	35.1a	37.06bcd
TGx 1988-18F	35.63abcde	37.68bc	34.39bcd	36.42ab	33.32a	35.49bcd
TGx 1988-22F	36.61bcdef	38.33bcd	35.95bcd	39.16b	35.02a	37.01bcd
TGx 1989-60F	36.39bcdef	39.03cd	34.22bcd	37.05ab	32.87a	35.91bcd
TGx 1990-129F	37.64cdef	38.88cd	36.84d	37.15ab	34.59a	37.02bcd
Magoye (check)	33.89ab	37.66bc	33.98bcd	36.06ab	33.73a	35.06abc
Safari (check)	34.66abc	36.88b	34.6bcd	35.22ab	32.38a	34.75ab
Lukanga (check)	32.56a	33.55a	29.61a	34.1a	35.65a	33.1a
Location Means	36.34b	38.23c	34.55a	36.9b	33.47a	35.90ab
Min	32.56	33.55	29.61	34.1	29.12	33.1
Max	39.29	40.12	37.02	39.16	35.65	37.57
Mean of IITA genotypes	37.00	38.77	35.01	37.34	33.36	36.30
Mean of checks	33.70	36.03	32.73	35.13	33.92	34.30
LSD 5%	1.0504	0.4375	1.359	1.756	4.702	1.297
CV	1.9	0.8	2.6	3.1	9	

Table 8: Mean Protein content (%) of fifteen soybean genotypes tested at the five different locations during the 2013/ 2014 season.

Simple correlation analysis of protein content with climatic and edaphic parameters (Appendix 2) showed simple relationships between protein and pH (r = 0.81), protein and calcium (r = 0.77), protein and Zinc (r = 0.69), protein and magnesium (r = 0.50), protein and manganese (r = 0.50), protein and iron (r = -0.97), Protein and rainfall (r = -0.93), protein and nitrogen (r = -0.75), protein and phosphorus (r = -0.47) and protein and sulphur (r = -0.45).

Stepwise regression analysis was done in order to determine the main causal factors to the fluctuations in protein content and results revealed that three edaphic factors namely, Phosphorus (P), Sulphur (S) and iron (Fe) were the most important causal factors with b-values of -0.055, -0.082 and -0.609 respectively and explaining up to 48% of the variation in oil content ($R^2 = 48.48$). The sensitivity of protein to change in phosphorus is implied by the magnitude of the associated b- values (b = -0.055). These results also showed that the other factors were not important determinants, to the observed variation in protein content.

The resulting regression equation for protein was;

Protein = 42.2 - 0.0548 P – 0.0825 S – 0.609 Fe R² adj = 47.8%

Combined analysis of variance for protein showed highly significant differences ($p \le 0.001$) among Locations and Genotypes. Significant Genotype by Location interactions were also observed (Table 9). This shows that the genotypes responded differently at the five locations implying that there was diversified genetic base for oil content for a breeder to select from.

Environments were significantly different from each other for protein content (Table 9) with GART having the highest location mean (38.23%) across all locations and Misamfu having the lowest location mean (33.47%) across all locations (Table 8). Similarly, significant differences among genotypes were evident when all fifteen genotypes were considered across the five locations with TGx 1830–20E a promiscuous (self nodulating) genotype having the highest genotype mean (37.57%) across locations and Lukanga (non- promiscuous genotype) having the lowest genotype mean (33.1%) for protein content across locations.

	Protein content (%)						
Source of variation	d.f.	(m.v.)	m.s.				
Location	4		107.303**				
Reps (L)	5		2.066				
Genotype	14		15.941**				
Location*Genotype	56		3.478*				
Residual	65	(5)	2.11				
Total	144	(5)					

Table 9: Combined analysis of variance across five locations for soybean protein content.

*, ** Significantly different at $p \le 0.05$ and $p \le 0.001$ levels respectively, d.f. = degree of freedom, m.s. = Mean Square

The results in table 8 showed that there was high interaction between genotypes and locations as the genotypes performance and ranking varied from one location to another. For instance, the best genotype at Kabwe TGx 1987-23F was third at GART and Msekera while being fourth and thirteenth at Masumba and Misamfu, respectively. Similarly, genotype TGx 1830-20E which was the second best performing genotype at Kabwe, fluctuated to fifth and first ranks at GART and Msekera, respectively. The genotype was second best performing at Masumba and sixth at Misamfu. The pattern of performance of Lukanga at Masumba was not easy to explain. On the other hand Safari showed similar fluctuations though generally having lower protein content.

4.3 Stability Analysis for Oil and Protein Content in Soybean

4.3.1 AMMI Model and Pattern Analysis for Soybean Oil Content

Source	Df	SS	MS	F	F_prob	Explained %
Treatments	74	261.84	3.538	11.51	< 0.001**	
Genotypes	14	71.08	5.077	16.52	< 0.001**	27.15
Environments	4	129.38	32.345	37.25	< 0.001**	49.41
Block	5	4.34	0.868	2.82	0.02276*	
Interactions	56	61.38	1.096	3.57	< 0.001**	23.44
IPCA1	17	40.44	2.379	7.74	< 0.001**	65.88
IPCA2	15	11.59	0.773	2.51	0.00544*	18.88
Residuals	24	9.35	0.39	1.27	0.22331	
Error	65	19.98	0.307			
Total	149	286.16	1.921			

Table 10: ANOVA for the AMMI Analysis of soybean oil content across five environments.

*, **: Significant at p<0.05 and p<0.001 level

The AMMI analysis of variance for soybean oil content of the fifteen genotypes tested in five environments showed highly significant differences (p<0.001) among genotype main effects, environment main effects and Genotype x environment interaction (Table 10). The model revealed that the differences between the environments accounted for 49.41% of the variation while the genotypes and the G x E interaction accounted for 27.15% and 23.44% of the variation respectively.

The IPCA1 and IPCA 2 parameters were significant at p < 0.001 and p < 0.05 respectively. The IPCA 1 accounted for 65.88% of the GE interaction sum of squares while IPCA 2 accounted for 18.88% of the variability. The two IPCA parameters explain 84.76% of the G x E sum of squares while the remaining 15.24% would be residual.



E= Environment, G= Genotype; E1= Kabwe, E2= GART, E3=Msekera, E4= Masumba, E5= Misamfu

Figure 3: AMMI 1 biplot of genotypes and environment IPCA 1 scores versus the soybean oil means of fifteen genotypes and five environments.

The AMMI 1 biplot for soybean oil content (figure 3) showed that the variability due to environments was higher than that due to genotypes as the points for environments were more scattered in the biplot than the points for genotypes.

Genotypes and environments with high stability and instability could be identified from figure 3. One check G15 (Lukanga) and one environment E5 (Misamfu) were dispersed away from the axes of the biplot showing high instability.

The IITA genotypes G11 (TGx 1989-60F), G1 (TGx 1740-2F), G12 (TGx 1990-129F) and the check G13 (Magoye) with IPCA1 scores close to zero and oil content close to the mean showed stability and general adaptability with negligible interaction. Genotype G11 (TGx 1989-60F) was the closest to the centre of the biplot exhibiting general adaptability and the best stability.

According to the AMMI 1 biplot in figure 3, the ideal genotypes that were stable with high oil content above the mean were G13 (Magoye) a check and the IITA genotypes G6 (TGx 1987-11F) and G8 (TGx 1988-9F) in quadrant II.

The biplot revealed that despite being widely adapted, some genotypes also had specific adaptability to some sites. These included genotypes such as G12 (TGx 1990-129F) and the check Magoye, which exhibited specific adaptability to E4 (Masumba), a high temperature and generally low rainfall environment (Table 1a and 1b). The other genotypes that had specific adaptability to E4 (Masumba) were G4 (TGx 1887-65F) and G6 (TGx 1987-23F).

The check G14 (Safari) and genotype G10 (TGx 1988-22F) showed specific adaptability for both E1 (Kabwe) and E3 (Msekera) which are medium rainfall environments and their interaction was positive as they fell in the same sector. No genotypes showed specific adaptability for E2 (GART).

The environments were spread around the biplot with the high soybean oil yielding environments in quadrant II and III and the lower soybean oil yielding environments in quadrants IV. The high oil potential environments falling on the right hand side of the midpoint of the main effect axis were E4 (Masumba) in quadrant II and E3 (Msekera) in quadrant III while environment E2 (GART) was a low oil yielding environment as it lay in quadrant IV.

Genotype	Mean Oil%	Oil % Rank	IPCA 1	IPCA 2
TGx 1740-2F	18.35	5	-0.0970	0.3236
TGx 1830-20E	16.73	15	0.7152	0.6160
TGx 1835-10E	17.25	13	0.2260	-0.2278
TGx 1887-65F	17.74	11	0.2379	0.3113
TGx 1904-6F	18.09	10	-0.3403	-0.2045
TGx 1987-11F	18.39	4	0.2203	0.0129
TGx 1987-23F	17.09	14	1.0468	0.4487
TGx 1988-9F	18.13	8.5	0.4425	-1.0111
TGx 1988-18F	18.13	8.5	-0.2185	-0.2351
TGx 1988-22F	18.63	3	-0.6472	-0.2927
TGx 1989-60F	18.27	7	-0.0263	0.0122
TGx 1990-129F	17.62	12	0.1178	-0.2546
Magoye (check)	18.33	6	0.1623	-0.2004
Safari (check)	18.95	2	-0.5931	0.2598
Lukanga (check)	19.47	1	-1.2462	0.4416
	Genotype TGx 1740-2F TGx 1830-20E TGx 1835-10E TGx 1835-10E TGx 1887-65F TGx 1904-6F TGx 1987-11F TGx 1987-23F TGx 1988-9F TGx 1988-18F TGx 1989-60F TGx 1990-129F Magoye (check) Safari (check) Lukanga (check)	GenotypeMean Oil%TGx 1740-2F18.35TGx 1830-20E16.73TGx 1835-10E17.25TGx 1835-10E17.25TGx 1887-65F17.74TGx 1904-6F18.09TGx 1987-11F18.39TGx 1987-23F17.09TGx 1988-9F18.13TGx 1988-18F18.13TGx 1988-22F18.63TGx 1989-60F18.27TGx 1990-129F17.62Magoye (check)18.33Safari (check)19.47	GenotypeMean Oil%Oil % RankTGx 1740-2F18.355TGx 1830-20E16.7315TGx 1835-10E17.2513TGx 1887-65F17.7411TGx 1904-6F18.0910TGx 1987-11F18.394TGx 1987-23F17.0914TGx 1988-9F18.138.5TGx 1988-18F18.138.5TGx 1988-22F18.633TGx 1989-60F18.277TGx 1990-129F17.6212Magoye (check)18.336Safari (check)18.952Lukanga (check)19.471	GenotypeMean Oil%Oil % RankIPCA 1TGx 1740-2F18.355-0.0970TGx 1830-20E16.73150.7152TGx 1835-10E17.25130.2260TGx 1887-65F17.74110.2379TGx 1904-6F18.0910-0.3403TGx 1987-11F18.3940.2203TGx 1987-23F17.09141.0468TGx 1988-9F18.138.50.4425TGx 1988-18F18.138.5-0.2185TGx 1988-22F18.633-0.6472TGx 1989-60F18.277-0.0263TGx 1990-129F17.62120.1178Magoye (check)18.3360.1623Safari (check)18.952-0.5931Lukanga (check)19.471-1.2462

Table 11. Mean Oil content and ranking order with IPCA 1 and IPCA 2 scores of fifteen genotypes tested across five environments.

AMMI predicted oil content means ranged from 17.09% to 19.47% across environments (Table 11). The check entries G15 (Lukanga) and G14 (safari) ranked first and second highest respectively with regards to mean oil content while the IITA genotype G2 (TGx 1830-20E) had the lowest mean oil content amongst the fifteen genotypes tested across the five environments.

The genotype IPCA1 scores for oil content ranged from -1.24621(G15- Lukanga) to 1.04677 (G7-TGx 1987-23F) with some genotypes having positive IPCA scores and some having negative IPCA scores (Table 11). It was observed that the genotypes that had the high oil content such as Lukanga also had high IPCA 1 scores (not close to zero) whether positive or negative. According to AMMI, genotypes with IPCA scores close to zero are considered to be stable across test locations and these were G11 (TGx 1989-60F) and G1 (TGx 1740-2F).

Environment Code	Environment	Oil mean %	IPCAe[1]	IPCAe[2]
E3	Msekera	18.98	-0.50129	0.15228
E5	Misamfu	18.86	1.77469	0.31922
E4	Masumba	18.13	0.17622	-0.75265
E1	Kabwe	18.03	-0.64082	-0.77487
E2	GART	16.38	-0.8088	1.05601

Table 12: Environment means and IPCA scores of the fifteen genotypes in the five environments

According to AMMI analysis (Table 12), E3 (Msekera) yielded the highest mean oil content (18.98%) among the five environments followed by E5 (Misamfu) which had 18.86% while E2 (GART) yielded the lowest oil content (16.38). E5 (Misamfu) had the highest IPCA 1 value (1.77469) among the environments tested and was therefore the most diverse environment. Masumba with an IPCA 1 score of 0.17622 was the most stable environment.



E= Environment, G= Genotype; E1= Kabwe, E2= GART, E3=Msekera, E4= Masumba, E5= Misamfu

Figure 4: AMMI 2 Biplot of IPCA1 scores versus the IPCA2 scores of fifteen genotypes and five environments for soybean oil content

When IPCA1 scores for oil content were plotted against their respective IPCA2 scores to further explore adaptation (figure 4) it was observed that, the IITA genotypes G11 (TGx 1989-60F), G6 (TGx 1987-11F), , G3 (TGx 1835-10E), G12 (TGx 1990-129F), G9 (TGx 1988-18F), G5 (TGx 1904-6F), , G1 (TGx 1740-2F), G4 (TGx 1887-65F) as well as the checks G13 (Magoye) and G14 (Safari) were scattered close to the centre of the Biplot. This was an indication of minimal environmental interaction and therefore stability of the genotypes. G11 (TGx 1989-60F), G6 (TGx 1987-11F) were closest to the centre of the biplot and therefore could be considered as the most stable genotypes for soybean oil content. The check G15 (Lukanga), G8 (TGx 1988-9F) and G7

(TGx 1987-23F) were furthest from the origin and therefore considered the most unstable genotypes for soybean oil content.

The highest interactions by environments were expressed by E5 (Misamfu), E2 (GART) and E1 (Kabwe) while E3 (Msekera) expressed the lowest interaction.

The biplot in figure 4 further showed that the genotypes G8 (TGx 1988-9F) had positive interaction with environment E4 (Masumba) hence exhibited specific adaptation to this environment.

The check G15 (Lukanga) had specific adaptation to E2 (GART) and E3 (Msekera) while G10 (TGx 1988-22F) showed specific adaptation to E1 (Kabwe). G2 (TGx 1830-20E) and G7 (TGx 1987-23F) had specific adaptation to E5 (Misamfu).

Code	Environment	Environment Mean	Score	1	2	3	4
E1	Kabwe	18.03	-0.6408	G15	G10	G14	G8
E2	GART	16.38	-0.8088	G15	G14	G10	G1
E3	Msekera	18.98	-0.5013	G15	G14	G10	G1
E4	Masumba	18.13	0.1762	G8	G15	G10	G14
E5	Misamfu	18.86	1.7747	G7	G6	G8	G13

 Table 13: The AMMI model's first four genotype selections for mean soybean oil content across five environments.

The best four performing genotypes in each environment as selected by AMMI are shown in Table 13. The Table indicates the best adapted genotypes in relation to the different environments with regards to soybean oil content. Two checks (G15 – Lukanga and G14 -Safari) and two IITA genotypes G10 (TGx 1988-18F), and G8 (TGx 1988-9F) were among the first four performing genotypes at four of the five environments tested. The check G15 (Lukanga) was the best

performing genotype at three of the five environments used in the study namely E1 (Kabwe), E2 (GART) and E3 (Msekera) while G8 (TGx1988-9F) and G7 (TGx 1987-23F) were the best performing genotypes at environments E4 (Masumba) and E5 (Misamfu) respectively.

4.3.1.1 The AMMI Stability Value for Oil Content

According to the AMMI stability value (ASV) (Table 14) of the second principal component, G11 (TGx 1989-60F) had the smallest ASV score and was therefore ranked as the most stable genotype across environments for oil content followed by G12 (TGx 1990-129F). G15 (Lukanga) had the highest ASV score suggesting unstable oil yield across environments. Generally, most of the IITA genotypes were more stable than the check entries apart from the check Magoye which ranked third.

Table 14. AMMI Stability Value (ASV) and ranking order with IPCA 1 and IPCA 2 scoresfor soybean oil content of fifteen genotypes tested across five environments.

Code	Genotype	IPCA 1	IPCA 2	ASV Value	ASV Rank
G1	TGx 1740-2F	-0.0970	0.3236	0.3709	4
G2	TGx 1830-20E	0.7152	0.6160	1.4710	13
G3	TGx 1835-10E	0.2260	-0.2278	0.4797	7
G4	TGx 1887-65F	0.2379	0.3113	0.5425	8
G5	TGx 1904-6F	-0.3403	-0.2045	0.6676	9
G6	TGx 1987-11F	0.2203	0.0129	0.4117	5
G7	TGx 1987-23F	1.0468	0.4487	2.0061	14
G8	TGx 1988-9F	0.4425	-1.0111	1.3060	12
G9	TGx 1988-18F	-0.2185	-0.2351	0.4709	6
G10	TGx 1988-22F	-0.6472	-0.2927	1.2439	11
G11	TGx 1989-60F	-0.0263	0.0122	0.0507	1
G12	TGx 1990-129F	0.1178	-0.2546	0.3365	2
G13	Magoye (check)	0.1623	-0.2004	0.3634	3
G14	Safari (check)	-0.5931	0.2598	1.1379	10
G15	Lukanga (check)	-1.2462	0.4416	2.3694	15

4.3.2 AMMI Model and Pattern Analysis for Soybean Protein Content

The AMMI analysis of variance for soybean protein content of the 15 genotypes tested in five environments (Table 15) showed that soybean protein content was significantly affected by environments, genotypes (p> 0.001) and genotype x environment interaction (p> 0.05) indicating the presence of genetic variation.

Source	Df	SS	MS	F	F_prob	Explained %
Treatments	74	847.1	11.45	5.43	< 0.001**	
Genotypes	14	223.2	15.94	7.55	< 0.001**	26.35
Environments	4	429.2	107.3	51.93	< 0.001**	50.67
Block	5	10.3	2.07	0.98	0.43722^{NS}	
Interactions	56	194.8	3.48	1.65	0.02631*	23.00
IPCA 1	17	121.1	7.13	3.38	< 0.001**	62.17
IPCA 2	15	36.9	2.46	1.17	0.31949 ^{NS}	18.94
Residuals	24	36.7	1.53	0.72	0.80863	
Error	65	137.2	2.11	*	*	
Total	149	994.6	6.68	*	*	

 Table 15: ANOVA for the AMMI Analysis of soybean protein content across five environments.

NS,*, **: Non Significant, Significant at p=0.05 and p=0.001 level

The AMMI analysis of variance for soybean protein content further showed that 50.67% of the treatment Sum of Squares (SS) was attributable to environmental effects, 26.35% to genotypic effects and 23%% to G x E interaction sum of squares effects (Table 15).

The AMMI model demonstrated the presence of G x E interactions, and this has been partitioned among the first and second IPCA (Interaction Principal Component Axes). The first principal component (IPCA1) was highly significant (p < 0.001) while the second principal component (IPCA 2) was not significant. The IPCA 1 and the IPCA 2 of the AMMI analysis accounted for 62.17% and 18.94% of the variability for soybean protein content respectively. These two IPCA parameters combined captured 81.11% of the G x E sum of squares. The remaining 18.89% was the residual.



G=Genotype, E= Environment

Figure 5. AMMI 1 Biplot of genotypes and environment IPCA1 scores versus the soybean protein means of fifteen genotypes and five environments.

The biplot in figure 5 shows that the points for environment are more scattered than the points for genotypes indicating that variability due to environments was higher than that due to genotypes differences which is in complete agreement with the ANOVA (Table 15). The biplot also showed one check G15 (Lukanga) and one environment E5 (Misamfu) dispersed away from the area of the biplot showing their large variability.

Three IITA genotypes namely G1 (TGx 1740 2F), G9 (TGx 1988-18F), G8 (TGx 1988-9F), and one check G14 (Safari) as well as one environment E4 (Masumba) were clustered near the center of the biplot indicating an average performance of the genotypes and environment.

IITA Genotypes G1 (TGx 1740-2F), G8 (TGx 1988-9F), G11 (TGx 1989-60F), G 12 (TGx 1990-129F) and the check G14 (Safari), with IPCA scores close or equal to zero and protein content close to the mean exhibited stability and general adaptability with negligible interaction. G1 (TGx 1740-2F) and G9 (TGx 1988-18F) were closest to the centre of the Biplot and were therefore the most stable genotypes.

The ideal genotypes which were stable and had high protein content were all IITA genotypes namely G8 (TGx 1988-9F), G12 (TGx 1990-129F), G11 (TGx 1989-60F), and G2 (TGx 1830-20E).

Although G8 (TGx 1988-9F), G11 (TGx 1989-60F) and G12 (TGx 1990-129F were widely adapted and stable genotypes, they formed a group around E4 (Masumba) environment indicating their adaptation to that environment which has low rainfall and high temperature.

The genotypes G2 (TGx 1740-2F) and G7 (TGx 1987-23F) had specific adaptability to the medium rainfall environments E2 (GART) and E1 (Kabwe) respectively because, despite having high mean performance, they had large IPCA values.

G3 (TGx 1835-10E) exhibited specific adaptability for E3 (Msekera) a medium rainfall environment with protein content less than the mean. Since this genotype and the environment had the same sign on the IPCA axis, their interaction was positive.

E1 (Kabwe), E2 (GART) and E4 (Masumba) falling on the right hand side of the midpoint of the main effect axis, seemed to be favorable environments for soybean protein content among the environments in the study. These were also high protein potential environments as they were found in quadrant II. The lower protein potential environments was E3 (Msekera) in quadrant I. The Biplot also indicated that E2 (GART) was the highest yielding environment as it was the furthest to the right of the midpoint.

Code	Genotype	Mean Protein%	Protein% Rank	IPCA 1	IPCA 2
G1	TGx 1740-2F	35.59	9	-0.0250	-0.3246
G2	TGx 1830-20E	37.57	1	0.5433	0.1024
G3	TGx 1835-10E	34.41	14	1.1145	-0.3939
G4	TGx 1887-65F	37.28	2	-0.1942	0.1947
G5	TGx 1904-6F	35.44	11	-0.4709	-0.4056
G6	TGx 1987-11F	35.8	8	0.4687	-1.2363
G7	TGx 1987-23F	36.98	6	1.1152	0.2633
G8	TGx 1988-9F	37.06	3	-0.0170	1.0961
G9	TGx 1988-18F	35.49	10	-0.1008	0.1169
G10	TGx 1988-22F	37.01	5	-0.2461	-0.2181
G11	TGx 1989-60F	35.91	7	0.2339	-0.2675
G12	TGx 1990-129F	37.02	4	0.0724	0.7191
G13	Magoye (check)	35.06	12	-0.4651	0.0438
G14	Safari (check)	34.75	13	0.0143	0.5580
G15	Lukanga (check)	33.09	15	-2.0431	-0.2484

 Table 16. Mean protein content and ranking order with IPCA 1 and IPCA 2 scores of fifteen

 genotypes tested across five environments.

AMMI predicted protein content means ranged from 33.09% to 37.57% across environments (Table 16). An IITA genotype (G2) TGx 1830-20E had the highest mean protein content followed by another IITA genotype (G4) TGx 1887-65F while a check (G15) Lukanga had the lowest protein content amongst the fifteen genotypes across the five environments.

The IPCA1 scores ranged from -0.46514 for G13 (Magoye) to 1.11445 G3 (TGx 1835-10E). Seven genotypes (G2, G3, G6, G7, G11, G12 and G14) had positive IPCA1 scores ranging from 0.01432 to 1.11523 while eight genotypes (G1, G4, G5, G8, G9, G10 G13 and G13) had negative IPCA 1 scores (Table 16). The check G15 (Lukanga) had the highest negative IPCA 1 score while G7 (TGx 1987-23F) had the highest positive IPCA 1 score.

Among the tested genotypes, G7 (TGx 1987-23F) an IITA genotype had the highest IPCA1 value of 1.11523 and therefore exhibiting specific adaptation, while G14 (Safari) a check and two IITA

genotypes G8 (TGx 1988-9F) and G1 (TGx 1740-2F) had IPCA1 scores closer to zero and can thus be considered to be more stable over all environments (Table 16).

Environ	ment code	Environment	Protein Mean %	IPCAe[1]	IPCAe[2]
]	E2	GART	38.23	0.63413	-0.40681
]	E4	Masumba	36.9	0.22904	-1.45832
]	E1	Kabwe	36.34	0.81895	0.22478
]	E3	Msekera	34.55	0.77797	1.37221
]	E5	Misamfu	33.47	-2.46009	0.26814

 Table 17. Environmental means and IPCA scores of the five environments for protein content (%)

AMMI predicted that E2 (GART) had the highest mean soybean protein content (38.23%) while E5 (Misamfu) had the lowest mean protein content (33.47%) (Table 17). Estimates for environmental IPCA 1 scores showed that Masumba with an IPCA 1 score of 0.22904 had the closest IPCA 1 score to zero and therefore was the most stable environment. Misamfu had the highest negative IPCA 1 score at -2.46009 and therefore was the most diverse environment for protein content.



Figure 6. AMMI 2 Biplot of IPCA1 scores versus the IPCA2 scores of fifteen genotypes and five environments for soybean protein content.

Figure 6 gives the AMMI 2 biplot, with the IPCA 1 and IPCA 2 for soybean protein content. Distribution of genotype points in the AMMI 2 biplot for protein content (Figure 6) revealed that six IITA genotypes namely G9 (TGx 1988-18F), G4 (TGx 1887-65F), G10 (TGx 1988-22F), G2 (TGx 1830-20E), G1 (TGx 1740-2F), G11 (TGx 1989-60F) and one check G13 (Magoye), were stable as they were close to the origin of the biplot indicating minimal interaction of these

genotypes with environments. The most stable genotypes for protein content were the IITA genotypes G9 (TGx 1988-18F), G4 (TGx 1887-65F) as their points were closest to the origin. The check G15 (Lukanga) and genotype G6 (TGx 1987-11F) were scattered farthest away from the origin of the biplot indicating that these genotypes were unstable in performance.

According to AMMI 2, the environments E1 (Kabwe) and E2 (GART) were the largest contributors to phenotypic stability of the fifteen genotypes as they were the closest to the biplot origin (Figure 6). On the other hand, E3 (Msekera), E5 (Misamfu) and E4 (Masumba) highly contributed to the GE interaction as they were positioned far from the biplot origin.

Genotypes G12 (TGx 1990-129F) and G8 (TGx 1988-9F) had positive interaction with environment E3 (Msekera), hence exhibited specific adaptation with this environment. G6 (TGx 1987-11F) had positive interaction with E4 (Masumba) while G3 (TGx 1835-10E) exhibited specific adaptation to E2 (GART).

Code	Environment	Envt. Mean	Score	1	2	3	4
E1	Kabwe	36.34	0.819	G2	G7	G8	G12
E2	GART	38.23	0.6341	G2	G7	G4	G10
E3	Msekera	34.55	0.778	G8	G7	G2	G12
E4	Masumba	36.9	0.229	G6	G2	G10	G4
E5	Misamfu	33.47	-2.4601	G15	G4	G10	G8

 Table 18: The AMMI model's first four genotype selections for mean soybean protein content across five environments.

Table 18 shows the best four AMMI genotype selections per environment with respect to soybean protein content. These results give an indication of the best adapted genotypes in relation to the different environments. Among the three checks used in the study, only one (Lukanga) made it to the top four highest protein yielding genotypes across environments. The IITA genotype G2 (TGx
1830-20E) was well adapted to four of the five environments tested but was best adapted to environment E1 (Kabwe) and E2 (GART). The genotypes G8 (TGx 1988-9F), G6 (TGx 1987-11F) and check G15 (Lukanga) were best adapted to environments E3 (Msekera), E4 (Masumba) and E5 (Misamfu) respectively. G7 (TGx 1987-23F) on the other hand performed well across three of the five environments tested namely E1 (Kabwe), E2 (GART) and E3 (Msekera).

4.3.2.1 The AMMI Stability Value for protein content

Based on the AMMI stability value (ASV) ranking (Table 19), G9 (TGx 1988-18F) an IITA genotype had the lowest ASV and therefore ranked first and was the most stable genotype for soybean protein content followed by G1 (TGx 1740- 2F) with G4 (TGx 1887-65F) ranking third. The genotype with the highest ASV and therefore lowest in rank and unstable was the check G15 (Lukanga).

Code	Genotype	IPCA 1	IPCA 2	ASV	Rank
G1	TGx 1740-2F	-0.0250	-0.3246	0.3278	2
G2	TGx 1830-20E	0.5433	0.1024	0.9896	10
G3	TGx 1835-10E	1.1145	-0.3939	2.0570	14
G4	TGx 1887-65F	-0.1942	0.1947	0.4021	3
G5	TGx 1904-6F	-0.4709	-0.4056	0.9447	9
G6	TGx 1987-11F	0.4687	-1.2363	1.4998	12
G7	TGx 1987-23F	1.1152	0.2633	2.0374	13
G8	TGx 1988-9F	-0.0170	1.0961	1.0966	11
G9	TGx 1988-18F	-0.1008	0.1169	0.2169	1
G10	TGx 1988-22F	-0.2461	-0.2181	0.4963	4
G11	TGx 1989-60F	0.2339	-0.2675	0.5010	5
G12	TGx 1990-129F	0.0724	0.7191	0.7310	7
G13	Magoye (check)	-0.4651	0.0438	0.8438	8
G14	Safari (check)	0.0143	0.5580	0.5586	6
G15	Lukanga (check)	-2.0431	-0.2484	3.7096	15

 Table 19. AMMI Stability Value (ASV) and ranking order with IPCA 1 and IPCA 2 scores

 for soybean Protein content of fifteen genotypes tested across five environments.

CHAPTER 5 DISCUSSION

5.1 Oil and Protein Content Variations of Soybean Genotypes

5.1.1 Oil Content of Soybean Genotypes

The mean oil content for the soybean genotypes tested in this study ranged from 16.73% to 19.47% across locations and is within the range that has been reported by other workers such as Ramana and Satyanarayana (2006) who in a study on soybean in India found that mean oil content ranged from 18.68% to 21. 63% with a general mean of 20.86%. Literature states that soybean is typically composed of approximately 20% oil (Padgette *et al.*, 1996; Carpenter *et al.*, 2002). The lower oil yields obtained from the current study could be attributed to specific and inherent genotypic expression (Rodrigues *et al.*, 2014; Brumm and Hurburgh, 2002; Bonato *et al.*, 2000), coupled with possible environmental influences given that these materials were not developed for Zambian conditions.

The results from correlation analysis were indicative of the importance of iron and rainfall in oil content of soybean but these relationships could not be concluded in the absence of a test for cause – effect relationship. Fluctuations in oil content among soybean genotypes evaluated were wide within locations and across locations. The within location differences can be attributed to inherent genetic differences among the genotypes. Information from the source of the genotypes, IITA, (personal communication) confirmed that the development of these materials did not focus on improvement of oil content hence the differences. Several studied (Gurmu *et al.*, 2009; Arslanoglu, 2011; Tubic *et al* 2011) have reported fluctuations in oil content among soybean genotypes.

This study revealed that oil content was different from location to location. The fact that the locations were different with regards to amount of rainfall, temperature and indeed soil types implied that the effect of these edaphic and climatic conditions were variable on the physiological processes in the synthesis of oil in soybean. The differential fluctuations in oil content from location to location observed in the current study manifesting genotype by environment interactions strongly suggested the influence of the environment in the expression of oil content trait. In this study, the environment was an aggregate of climatic and edaphic factors. From the current study results, the key environmental factors important for changes in oil content were iron

(Fe) and rainfall, explaining 48% of the variation in the oil content. The oil content in soybean is influenced by many factors both edaphic and climatic (Piper and Boote, 1999; Arslanoglu, 2011; Silva *et al* 2016). Several researchers have reported that the oil content of soybean seed varies with environmental conditions encountered during the growing period, such as temperature and rain (Rose, 1988; Gurmu *et al.*, 2009; Rodrigues *et al.*, 2014). Further, Bellaloui *et al* 2015 stated that oil biosynthesis is dependent on enzyme activity, the nutrients that are absorbed from soil which contribute to enzyme activity and other environmental factors

Dornbos and Mullen, (1992) reported that severe drought stress during seed filling stage could result in up to 12.4% oil decrease in soybean seed. Further, Rotundo and Westgate,(2009) in a meta- analysis of published data quantifying the effect of water and temperature stress on soybean seed oil and protein accumulation also concluded that water stress can reduce oil content by about 25%, but the timing of stress was also very important. When water stress is experienced early, in the reproductive period, the magnitude of the effect size was not different from zero while water stress during seed filling decreased oil content by about 35%. In the current study, the lowest mean rainfall was at Kabwe and the oil content ranked fourth (18.03 %) compared to most of wetter locations such as Misamfu and Msekera where the oil content was above 18% apart from GART. Rotundo and Westgate, (2009) and Bellaloui *et al.*, (2013) also found that drought/ water stress decreased soybean oil content. Yamagata *et al.*, 1987 stated that oil and carbohydrate synthesis by the seed, is primarily dependent on concurrent carbon fixation during the seed filling stage. A reduction in assimilate supply due to water stress during seed fill could therefore directly impact on the synthesis of oil and residual components.

With regards to Iron (Fe), the current study revealed that the soils at Misamfu had the highest iron (10.23 mg/kg) while soils at GART had the lowest iron (3.38 mg/kg). Results also showed that the second highest oil content among the locations tested was realized from Misamfu (18.86%) while the lowest was from GART (16.38%). Kobraee and Shamsi, (2015) reported that availability of iron increased oil contents however, excess amounts of iron in the soil reduced oil content of soybean. Iron is an important element for synthesis of chlorophyll, metabolism and is also used in many plant enzyme systems. Iron is also required for respiration, DNA synthesis and photosynthesis (Burton, 1998). Iron deficiency hampers chlorophyll production which in turn affects plant growth and could result in death of the plant. The low levels of iron in the soils at

GART could therefore have contributed to the low oil content at this site as a result of hampered functioning of plant enzyme systems for important plant processes as well as reduced photosynthesis which produces the much needed Carbon, Hydrogen and Oxygen for fatty acid synthesis.

Genotypes were found to be different for their oil content in the current study and given the differences in their genetic constitutions, some being improved varieties' while others were less so, and their inherent ability to synthesize oil was manifested. Lukanga, a check entry, had the highest mean oil content (19.47%) among all the tested genotypes while TGx 1830-20E had the lowest mean oil content (16.73%) across locations. Genotypic differences in oil content in soybean have been widely documented (Wilcox, 1985; Burton, 1989; Bonato *et al.*, 2000; Jaureguy *et al.*, 2011; Rodrigues *et al.*, 2014). The amount and composition (individual fatty acids) of oil, in soybean is known to be affected by genetics and the environment, specifically drought and temperature, as well as their interactions (Bellaloui *et al*, 2013). The significant differences among genotypes for soybean oil content in the ANOVA suggested that there was variability for this variable among the genotypes thus selection for superior ones was possible.

These differences, however, were not consistent as there were differential genotypic responses observed as evidenced through the significant genotype by location interactions. Differential responses implied that each given genotype had different reaction to existing edaphic and climatic conditions at each location, some being more consistent and others less so. Substantial differences in genotypic response across environments were observed for oil content as the genotypes changed in both magnitude of oil and protein content yield as well as ranking. For instance, TGx 1988-9F fluctuated from being the best genotype for oil content at Masumba to being fourth, tenth, thirteenth and eleventh at Misamfu, Kabwe, Gart and Msekera, respectively. The change in magnitude or ranking of a genotype from one location to another is an indication of significant genotype by environment (G x E) interactions. Kaya *et al*, 2002 reported that depending upon the magnitude of the interactions or the differential genotypic responses to environments, the varietal rankings can differ greatly across environments. Higher magnitude of change in genotypic means due to environments indicates considerable differences between environments for a given variable and also indicates that the variable was greatly influenced by environments. This therefore

suggests that differences between environments along with a large part of genotypic response is a direct result of environments (Tyagi and Khan, 2010).

The range of performance for oil content in this study was highest at GART from 19.3% to 15.05% i.e (4.25%) and lowest at Misamfu from 19.7% to 18.09% i.e. (1.61%). This means that the effect of the environments on oil content was higher at GART and less so at Misamfu. Singh and Chaudhary (1985b) in a study on soybean in India also found significant genotypic differences and genotype by environment interactions for oil content. Similar results were obtained by Rocha *et al.* (2002), Pfeiffer *et al.* (1995) and Bueno *et al;* (2013) and who found significant genotype by environment interactions for soybean oil content.

From the above, it is clear that genotypes' performance with regard to oil content is inherently different (genotypic) and also that these differences are differential with regard to the environment under which genotypes are grown. The significant GE interaction reduces the reliability of the mean performance across all environments as an indicator of superiority. To this extent therefore, additional decomposition of the GE Interaction is important.

5.1.2 **Protein Content of Soybean Genotypes**

With regards to protein content, the mean protein content for the genotypes tested in this study ranged from 33.09% (Lukanga) to 37.57% (TGx 1830-20E) with an overall mean of 35.9% across locations and are within the ranges reported by other workers. Arslanoglu *et al.*, (2011) who in a study conducted in Turkey found that protein content ranged from 29.25% to 38.57%. The differences in the genotypes' protein contents could be attributed to inherent genotypic expression (Rotundo and Westgate, 2009; Popovic *et al.*, 2013). The high protein content for TGx 1830- 20E could be attributed to TGx 1830- 20E's ability to fix atmospheric nitrogen without requiring inoculation unlike Lukanga, a non- promiscuous genotype that requires suitable rhizobia strains in order to fix nitrogen from the air (Tefera, 2011). Symbiotic nitrogen fixation is the main source of nitrogen in legumes such as soybean and is regarded as the main factor for seed protein content (Fabre and Plancho, 2000). Leffel et al., 1992 stated that improved nitrogen fixation could facilitate high seed protein content in soybean. Further, Maphosa, 2015 in a study on effect of inoculation

of soybean on nutritional quality parameters found that seed of promiscuous varieties contained significantly higher crude protein as compared to non-promiscuous varieties.

The results from correlation analysis indicated the importance of pH, calcium, iron, phosphorus, sulphur and rainfall in protein content of soybean but these relationships could not be concluded in the absence of a test for cause – effect relationship.

Results from the current study also revealed that protein content significantly differed within locations and from one location to another. The differences within locations can be attributed to inherent genetic differences among the genotypes as development of these lines did not specifically focus on improvement of protein content resulting in the observed differences. Bueno *et al.*, 2013 stated that genotypic variance is considered one of the most important parameters for quantifying the breeding potential of a population, and the existence of genotypic variance among the tested averages indicates the viability of the use of selective techniques in genotypes. Genotypic variation for seed protein concentration in soybean has been studied and documented (Thorne and Fehr, 1970; Wehrmann *et al.*, 1987; Wilcox and Cavins, 1995; Rodrigues *et al.*, 2014).

The differences in the protein contents at the different locations can also be attributed to the effect of differences in the environmental (edaphic and climatic) conditions at the various locations on accumulation of protein in soybean seeds. From the current study results, the key environmental factors important for changes in protein content were iron (Fe), Sulphur (S) and Phosphorus (P) explaining 48% of the variation in the protein content. The protein content in soybean is influenced by many factors both edaphic and climatic (Gibson and Mullen, 2001; Bellaloui, 2015). Several studies have reported that environmental conditions have great effect on protein content of soybean seed (Fehr *et al.*, 2003; Ning *et al.*, 2003; Zhang *et al.*, 2005).

In the current study, Misamfu had the highest levels of sulphur in the soil (23.18mg/kg) and yielded the lowest protein content (33.47%) among the locations tested. On the other hand, Masumba had the lowest soil sulphur (12.82%) levels and the second highest protein content (36.9%) was realized from this site. Contrary to this, Devi *et al.*, 2012 in a study conducted in India found that soybeans grown on soils with high levels of sulphur had high protein content. This result was

attributed to involvement of sulphur in increasing quality of protein through synthesis of amino acids such as cysteine and methionine. Similar results were reported by Havlin *et al.*, 1999. The difference in results obtained from the current study could be attributed to the fact that sulphur is mobile in soils and can easily be lost by leaching (Hellal and Abdelhamid, 2013) and could have therefore been leached as a result of high rainfall in Misamfu reducing the amount of sulphur during seed filling period as compared to the levels at the start of the season when soil samples were collected and parameters measured.

With regards to phosphorus, results from the current study revealed that Kabwe had the highest levels of phosphorus (P) in the soil (15.21 mg/kg) and protein content (36.34%) at this site ranked third. Masumba on the other hand had the lowest phosphorus level (1.99 mg/kg) and protein (36.99%) at this site ranked second. Similar results were obtained by Win *et al.*, 2010 who in a study carried out in Thailand observed that, soybean protein content decreased with increased P rates as very high soil phosphate values may have depressed the seed protein content. Phosphorus enables plants to convert solar energy into chemical energy, and plants need chemical energy to synthesize sugars, starches, and proteins. Phosphorus is relatively immobile in soils and the reason could be due to the fact that Phosphorus reacts in the soil with iron, aluminium and calcium and becomes unavailable for plant use (Hellal and Abdelhamid, 2013). Phosphorus deficiency impairs nodulation by affecting the assembly of functional iron- sulphur cluster (Burton, *et al.*, 1998) and this could have also been the reason for the reduced protein levels at Kabwe.

On the other hand, results from the study indicated that soils at Misamfu had the highest levels of iron (Fe) and the lowest protein content was obtained at this site while GART had the lowest levels of iron and the highest protein content. Kobraee and Shamsi, 2015 also found that excess amount of iron reduced protein content. Further, the reduced protein content of genotypes at this site could have been as a result of the known inverse relationship between protein and oil (Schwender *et al.*, 2003; Popovic *et al.*, 2013) as genotypes at this site yielded the highest oil content.

In addition to genetic variability, presence of significant genotype by environment interactions as revealed by the combined ANOVA were confirmed by the differential responses of the genotypes to environments causing the soybean protein content of the genotypes to vary from location to location, with some locations being suitable for harvesting of either high or low soybean protein levels owing to the varying edaphic and climatic conditions at each location. The range of performance for protein content was highest at Msekera where the range was from 37.02% to 29.61% i.e. 7.42% and lowest at Masumba from 39.16% to 34.1% i.e. 5.06% implying greater environmental effect on protein content at Msekera and minimal effect at Masumba. Other workers such as Gibson and Mullen (1996); Isaza (2002); Kumar *et al.* (2006) and Arslanoglu *et al* (2011) and also studied the effect of growing environment and genotype by environment interactions on oil and protein content of soybean genotypes grown at various locations and found significant genotype by environment effects on the studied traits.

Kang and Gorman (1989) stated that genotype by environment (G x E) interactions greatly reduce the significance of the correlation between phenotypic and genotypic values. When the interaction is due to variation caused by unpredictable environmental factors (e.g. rainfall or temperature) the breeder should develop widely adaptable varieties. These conclusions could be applicable the protein combined ANOVA results obtained in the current study which revealed significant differences among genotypes, environments and genotype by environment interaction. The significant genotype by environment interactions suggested that testing the selected genotypes under the given environments was important for genotypic comparisons and to determine the phenotypic stability of each genotype.

As was observed above under oil content, performance of the genotypes with regard to protein content also showed inherent genotypic differences and the response of the genotypes was different both within and across locations. Significant GE Interactions renders mean performance less useful in selection of the best performing genotypes as genotypes' relative ranking vary across the environments resulting therefore in the need for further decomposition of the GE Interaction.

5.2 Stability of Oil Content and Protein Content across Varying Environments

Clearly, every genotype has a set of environments that are best suitable for them, with respect to particular characteristic(s), but the implication of specific adaptation presents challenges in crop variety development and deployment as such, the concept of stability should be employed to enhance crop productivity in the wake of unpredictability of climate.

The significant genotype by environment interactions observed in this study and explained above has implications in identification of environments suitable for increased productivity with regard to oil and protein content of soybeans: That environments ranked genotypes differently implies that for each genotype a specific environment was most suitable.

In the present study, both the ANOVA for AMMI analysis of soybean oil content and protein content revealed that there was variability in the main effects (genotypes and environments) as well as the interaction effect (genotype by environment interaction). The significant differences (p < 0.05) between environments, genotypes and genotype by environment interaction indicated the presence of genetic variation for the two variables and the possibility of selecting of stable entries. Gurmu *et al* (2009) also made similar conclusions after studying the effect of genotype by environment interactions and stability of twenty soybean genotypes at six environments.

The soybean oil and protein AMMI ANOVA models both revealed that the differences between the environments accounted for a larger part of the variation followed by genotypes and the genotype by environment interaction respectively. This partitioning of the treatment sum of squares indicates that the environment effect was a predominant source of variation for the two variables followed by the genotype effect and GE interaction effect respectively. The significant environmental effect further suggested that the differential performance of the genotypes across the different environments could be explained by the fluctuations in the climatic and environment variables from one environment to another.

The large sums of squares for environment in the two AMMI models (49.41% and 50.67% for oil and protein content respectively) also indicates that the environments were diverse with large differences among environmental means causing most of the variation in soybean oil and protein content for the genotypes tested. Several previous studies reported that environmental conditions had the greatest effect on the oil and protein content of soybean seeds (Fehr *et al.*, 2003; *Zhang et al.*, 2005).

Both the AMMI 1 biplot for soybean oil content (Figure 2) and the AMMI biplot 1 for protein content (Figure 4) showed that the variability due to environments was higher than that due to genotypes as the points for environments were more scattered in the biplots than the points for

genotypes. This is an indication that there was good variation in the environments sampled and is in complete agreement with the results obtained in the AMMI ANOVA for both variables in Table 12 and Table 17 respectively. Kang, 2002 and Marjanovic- Jeromela *et al*, 2011 stated that genotype stability is considered a reaction to changing environmental conditions, which depend on unpredictable variation components. In this study, climatic conditions were the source of this variation component.

Genotype G11 (TGx 1989-60F) was the closest to the centre of the biplot showing the best stability for oil content followed by G1 (TGx 1740-2F) and G12 (TGx 1990- 129F) Figure (2). For protein content on the other hand, G1 (TGx 1740-2F) and G9 (TGx 1988- 18F) were closest to the centre of the Biplot and were therefore the most stable genotypes adaptable to a wide range of growing areas (Figure 4). This observed response agreed with Chapman *et* al (1997) who suggested that genotypes with lower IPCA scores and clustered near zero or the centre are said to be stable.

According to the AMMI model, genotypes with means greater than the grand mean and with IPCA score close to zero are considered as having general adaptation to all environments. However, genotypes with high mean performance and large IPCA score are considered as being specifically adapted to certain environments. Therefore, the generally adapted (ideal) genotypes that were stable with high oil content above the mean were G13 (Magoye), G6 (TGx 1987-11F) and G8 (TGx 1988-9F) in quadrant II of the biplot. For protein content on the other hand, G8 (TGx 1988-9F), G12 (TGx 1990-129F), G11 (TGx 1989-60F), and G2 (TGx 1830-20E) were the stable high protein yielding genotypes. Rashidi *et al.* (2013) stated that on a biplot, the points for the generally adapted genotypes would be at the right hand side of grand mean levels (suggesting high mean performance) and close to the line showing IPCA= 0 and (suggesting negligible or no $G \times E$ Interaction). Pacheco et al. (2005) also stated that for cultivar recommendation purposes, stable genotypes should also have desirable characteristics. This agrees with Ebehart and Russel (1966) as well as Nkhoma (2013) who recommended that breeders aim at developing varieties that are not only stable but also have above average performance in other traits. This means the named genotypes can be selected for breeding for high oil and protein content respectively in all the five environments. In other words, these genotypes can be recommended for wider adaptation and for production of high oil/ protein content in soybean.

Further, according to the AMMI Biplot analysis, regardless of their IPCA 1 scores direction/ sign, environments such as E1 (Kabwe), E3 (Msekera) and E4 (Masumba) falling on the right hand side of the midpoint of the main effect axis (Figure 2) were regarded to be favorable environments for soybean oil content among the genotypes evaluated while in Figure 4, E1 (Kabwe), E2 (GART) and E4 (Masumba) were regarded to be favorable environments for soybean protein content.

In this study, AMMI analysis also revealed that genotypes with very high IPCA scores were more specifically adapted to certain environments/ locations where they could fully exploit their potential (Table 13 and Table 18) while those with low IPCA scores were stable over a wide range of environments. This agrees with the findings of Zobel *et al.* (1988) and Crossa *et al.* (1997) who reported that the greater the IPCA score, whether negative or positive, the more specifically adapted the genotypes are to certain agro- ecology environments. The more the IPCA scores approximate zero (0), the more stable the genotype is over all the environments sampled. AMMI ranked (G15) Lukanga (IPCA1 -1.2462) as the highest oil yielder across environments while (G2) TGx 1830- 20E had the lowest oil content (Table 13). On the other hand, (G2) TGx 1830-20E (IPCA1 0.5433) was ranked as the highest protein yielder across environments while (G15) Lukanga was the lowest protein yielder. This shows that the genotype with the highest oil content yielded the lowest protein content across locations which agrees with results obtained in Table 5 and Table 8.

It was observed that the genotypes that yielded the highest oil and protein content also had high IPCA scores as compared to the other genotypes, an indication that they were not stable (Table 13 and Table 18). Pacheco *et al.* (2005) stated that selection for better stability generally results in lower mean yields and, conversely, that selection for higher mean yields may lead to poorer stability. Abalo *et al.*, (2003) and Asio, (2004) similarly reported that yield stability could only be expected from low yielding genotypes which do not exploit favorable environments. The implication therefore is that the more stable the genotype is, the lower yielding it becomes.

The AMMI analysis revealed significant environmental effect on the oil and protein contents of the tested genotypes from one location to another (Table 14 and Table 19). E3 (Msekera) yielded the highest mean oil content (18.98%) while E2 (GART) yielded the lowest (16.38%) mean oil content among the environments used in the study. Results showed that Masumba with an IPCA 1 score of 0.17622 exhibited minimum interaction effect of climatic conditions and could thus be

said to be a favorable environment for evaluating the performance of soybean genotypes for oil content.

AMMI predicted that E2 (GART) had the highest mean soybean protein content (38.23%) while E5 (Misamfu) had the lowest mean protein content (33.47%). Estimates for environmental IPCA 1 scores showed that Masumba with an IPCA 1 score of 0.22904 had the closest IPCA 1 score to zero and therefore was the most stable. Misamfu had the highest negative IPCA 1 score at -2.46009 and therefore was the most diverse environment.

Generally, oil content was observed to increase in environments that produced lower protein content. The results from the AMMI analysis in table 14 are in agreement with the results obtained from the analysis of variance in Table 5 and Table 8. These results are also similar to the findings of Miladinovic *et al*, (2006) and Arslanoglu *et al*., (2011) who found that, protein content generally increased in environments that produced lower oil content. The results are also in agreement with the findings of Tubic *et al*. (2011) who stated that besides the influence of genotypes, significant influence was also exerted by environmental factors on the negative correlation of these two soybean seed constituents.

When IPCA 1 scores were plotted against IPCA 2 scores (Figure 3 and Figure 5) to further explore adaptation, the AMMI 2 Biplot analysis showed that some genotypes had general adaptation while some had specific adaptability. The AMMI 2 biplot diagrams present the interaction effect. Marjanovic-Jeromela *et al.* (2011) indicated that the differences and genotype distributions in the Biplot are a consequence of genotype variations in different conditions/ environments. Figure 3 showed that the genotypes G8 (TGx 1988-9F), G15 (Lukanga) and G7 (TGx 1987-23F) expressed the highest interaction for soybean oil content indicating their narrow adaptability to certain environments and high sensitivity to environmental interactive forces while the genotypes G6 (TGx 1987-11F), G11 (TGx 1989-60F were the closest to the centre of the biplot expressing the lowest interaction with environments and indicating their stability or broad adaptability. Generally, environments with scores near zero have little interaction across genotypes and provide low discrimination among genotypes (Anandan *et al.*, 2009; Marjanovic -Jeromela *et al.*, 2011,)

For soybean protein content on the other hand, genotypes G15 (Lukanga) and G6 (TGx 1987-11F) were unstable in performance while G9 (TGx 1988-18F) and G4 (TGx 1887-65F) were the most stable genotypes for soybean protein content (Figure 5).

The largest environmental contributor to phenotypic stability of oil content for the tested genotypes was E3 (Msekera) Figure (3) while the largest environmental contributor to phenotypic stability of protein content were E1 (Kabwe) and E2 (GART) for protein content Figure (5) as they were the closest to the Biplot origin. This means that these environments were the most suitable for production of soybean genotypes with high oil and protein content respectively. The environments that highly contributed to the G x E interaction were E5 (Misamfu), E2 (GART) and E1 (Kabwe) for soybean oil content (Figure 3) and E3 (Msekera), E5 (Misamfu) and E4 (Masumba) for protein content (Figure 5) as they were positioned furthest from the Biplot origin.

The genotypes G8 (TGx 1988-9F), G15 (Lukanga), G10 (TGx 1988-22F) had positive interaction with the environments E4 (Masumba), E2 (GART) and E3 (Msekera), and E1 (Kabwe) respectively (Figure 3) as they fell in the same sector/ or exhibited acute angles with these environments. On the other hand, G2 (TGx 1830-20E) and G7 (TGx 1987-23F) had specific adaptation to the high rainfall environment E5 (Misamfu). This means that these genotypes were best suited or specifically adapted to the climatic conditions in these areas for soybean oil production. The smaller the angle between the interaction vectors is, the greater the similarity in the interaction response. Rashidi et al. (2013) stated that the AMMI analysis permits estimation of interaction effect of a genotype in each environment and it helps to identify genotypes best suited for specific environmental conditions. Genotypes and environments that fall into the same sector interact positively while those that fall into opposite sectors interact negatively (Osiru et al., 2009). A genotype showing high positive interaction with a given environment obviously has the ability to exploit the agro-ecological or agro-management conditions of the specific environment and is therefore best suited to that environment. Genotypes and environments positioned close to each other in the Biplot have positive associations, thus these enable the creation of agronomic zones with relative ease (Silveira et al., 2013). The interaction effect of a genotype in a given environment is approximated by projecting the genotype point onto the line determined by the environmental vector, where distance from the origin provides information about the magnitude of the interaction. The angle between the vectors of the genotype and environment is an indication of the nature of the interaction i.e. the interaction is positive for acute angles, negligible for right angles, and negative for obtuse angles.

Ranking of the genotypes in terms of stability using the ASV as suggested by Purchase (1997) showed that for oil content, G11 (TGx 1989-60F) had the least ASV score and was therefore the most stable genotype for soybean oil content across environments. G15 (Lukanga) had the highest ASV implying that its oil content was unstable across environments and suggesting specified adaptation to certain environments. These results are in agreement with results from the AMMI 1 (Figure 2) and AMMI 2 models for oil content.

For protein content, G9 (TGx 1988-18F) was ranked as the most stable genotype across environments owing to its smaller ASV while G15 (Lukanga) was more specifically adapted to certain environments as it had the largest ASV. This agrees with the results from the AMMI 2 model (Figure 5). These results however slightly differ from those obtained from the AMMI 1 model (Figure 4) where G1 (TGx 1740-2F) was identified as the most stable genotype followed by G9 (TGx 1988-18F). G15 (Lukanga) on the other hand maintained its rank as the most unstable genotype. According to Peterson *et al.* (1992), the genotype stability definition and response for quality parameters is relatively different from that conventionally used to characterize yield stability. For breeders, stability of quality properties is important from the points of changing genotypes ranks' throughout environments and influences selection efficiency.

In Table 15 and Table 20, the AMMI model successfully summarizes the patterns and relationships of genotypes and environments by showing the best four performing genotypes at each location for soybean oil and protein content respectively. For instance, for soybean oil content, G15 (Lukanga) was the best performing genotype in environments E1 (Kabwe), E2 (GART) and E3 (Msekera). This is an indication of the AMMI model's ability to analyse the GEI and identification of superior genotypes. From these results, it is evident the AMMI model can also be used in the selection of the most suitable environments for production and/ or evaluation of specific genotypes.

CHAPTER 6 CONCLUSIONS

This study revealed that genotypes were different for oil content and protein content and these differences were inherent (genotypic). The differences were however not consistent due to significant environmental influence on oil and protein content of soybean manifested through significant GE interactions. There was variation in the ranking of the genotypes within individual locations for oil content and protein content which made it difficult to identify superior genotypes.

Msekera was the most suitable environment for production of soybean genotypes with high oil content while GART was the most suitable environment for production of soybean genotypes with high protein content as these environments were the largest contributors to phenotypic stability of the named traits respectively. The two environments can also be recommended for use in the improvement of the two traits.

The genotypes Lukanga, Safari, TGx 1988-22F and TGx 1740-2F were best suited for Msekera with regard to oil content while genotypes TGx 1830-20E, TGx 1987-23F, TGx 1887-65F and TGx 1988-22F were best suited for protein content at GART. These genotypes can therefore be said to be best adapted to these environments and could be deployed to these areas as they would fully exploit their potential in the named environments.

The study further revealed that genotype TGx 1989-60F was the most stable variety for soybean oil, adaptable to a wide range of growing areas and may be suitable as a parental line in crosses to improve soybean for oil stability.

Genotypes TGx 1740-2F, TGx 1988-18F, TGx 1988-9F and Magoye were the stable genotypes for protein content. The AMMI Stability Value (ASV) Ranking for Protein indicated that TGx 1988-18F was the most stable genotype. These genotypes which were found to be stable and well adopted to all environments would be useful for exploitation as elite gene pool materials in future breeding programmes or for commercial exploitation.

Ideal genotypes were Magoye, TGx 1987-11F and TGx 1988-9F for oil content and G8 (TGx 1988-9F), G12 (TGx 1990-129F), G11 (TGx 1989-60F), and G2 (TGx 1830-20E) for protein content. These genotypes can therefore be recommended for wider adaptation and for production

of high oil and protein content respectively in soybean across different environments. The processing industry may also find interest in growing this genotype on larger production areas, for above average oil contents, and the overall stability of the grain oil composition.

It is important to note that this study was only conducted in one season and the results indicated significant G x E effects for protein and oil. Therefore, it is recommended that the study be repeated to assess the year effects at a location. The information is important to guide breeders and agronomists in variety development.

CHAPTER 7 REFERENCES

Abalo, G., J. J. Hakiza, R. El-Bedewy and E. Adipala. 2003. Genotype x Environment interactions studies on yield of selected potato genotypes in Uganda. African Crop Journal 11: 9-15

Acquaah, G. 2005. Principles of Crop production: theory, techniques, and technology. 2nd ed. Prentice Hall.

Alberts, M. 2004. A comparison of statistical methods to describe genotype environment interaction and yield stability in multi-location maize trials. Msc Thesis, department of plant science (plant breeding), Faculty of Agriculture., Univ. of Orange Free State, Bloemfontein, South Africa

Allard, R.W. 1964. Implications of genotype-environment interactions in applied plant breeding. Crop Sci., 4: 503-507.

Anandan, A., T. Sabesan, R. Eswaran, G. Rajiv, N. Muthalagan and R.Suresh. 2009. Appraisal of environmental interaction on quality traits of rice by additive main effects and multiplicative interaction analysis. Cereal research Communication, V37 (1), p.131-140

Annicchiarico, P. 1997. Joint regression vs. AMMI analysis of genotypes environment interactions for cereals in Italy. Euphytica *94*, *53-62*.

Aparicio, A. Redondo Cuenca, M. J. Villanueva-Suárez and M. A. Zapata-Revilla.Soybean, a promising health source, Nutricion Hospitalaria, 2008; Vol 23(4): pp 305-312.

Arioglu, H.H. 2007. Breeding and growing of oil plants. (Lesson books) Publications of Cukurova University, no: 220, Adana-Turkey.

Arslanoglu, F., A.S. Selim and E.K. Oner. 2011. Effect of genotype and environment interaction on oil and protein content of soybean (*Glycine max* (L.) Merrill) seed, African Journal of Biotechnology Vol. 10(80), pp. 18409-18417.

Asio, M. T. 2004. Agronomicc evaluation of some local and improved cowpea (*Vigna unguiculata*) lines in Uganda. Msc Thesis, Makerere University, Kampala, Uganda.

Baianu, I.C, T. You, D.M. Costescu, P.R. Lozano, V. Prisecaru, and R.L Nelson. 2010 Determination of Soybean Oil, Protein and Amino Acid Residues in Soybean Seeds by High Resolution Nuclear Magnetic Resonance (NMRS) and Near Infrared (NIRS). Nature Precedings. Becker, H.C, J. Leon (1988) Stability analysis in plant breeding. Plant Breeding 101(1): 1-23.

Bellaloui, N., Mengistu A., Kassem M. A., 2013. Effects of Genetics and Environment on Fatty Acid Stability in Soybean Seed. Food and Nutrition Sciences, 2013, 4, 165-175

Bellaloui, N., <u>H. A.</u> Bruns, <u>H.</u> K. Abbas, <u>A.</u> Mengistu, <u>D.</u> K. Fisher, and <u>K.</u> N. Reddy. 2015. Agricultural practices altered soybean seed protein, oil, fatty acids, sugars, and minerals in the Midsouth USA. Front Plant Sci. 6:31

Bennett. J.O. 2005, Soybean Seed Components as Affected by Nodal Position, Environmental Conditions, and Irrigation. MSc. Thesis, University of Missouri-Columbia.

Berger, J.D., J. Speijers, R.L. Sapra and U.C. Sud. 2007. Genotype by environment interaction and chickpea improvement, In: Chickpea Breeding and Management. (Eds), S.S. Yadav, R. Redden, W. Chen, B. Sharma. pp. 617-629 (CABI: Wallingford, UK).

Blanusa, T., R. Stikic, B. Vucelic-Radovic, M. Barac, D. Velickovic. 2000. Dynamics of seed protein biosynthesis in two soybean genotypes differing in drought susceptibility. Biologia Plantarum *43*: 55-59.

Bonato, E. R., P.F. Bertagnolli, C.E.Lange, S.A.L. Rubin, 2000. Oil and Protein Content in Soybean Genotypes Developed after 1990. Pesq. Agro. Bras. 35(12):2391-2398.

Brumm, T.J., and C.R. Hurburgh .2002. Quality of the 2002 soybean crop from the United States. American Soybean Association Quality Mission to Asia, December 13, 2002.

Boucher, D. 2011. Soybeans In: The Root of the Problem; What's driving tropical deforestation today? Union of concerned scientists.

Bueno, R.D., L. L. Borges, K. M. A. Arruda, L. L. Bhering, E.G. Barros, M. A. Moreira. 2013. Genetic parameters and genotype x environment interaction for productivity, oil and protein content in soybean. African Journal of Agricultural Research. 8(38) pp. 4853-4859.

Bunyolo, A., B. Chirwa, and M. Muchinda. 1997. Agro-ecological and climatic conditions.

p. 19-27. In: S.W. Muliokela (Ed.) Zambia seed technology handbook. Berlings, Arlov, Sweden.

Burton, J.W. 1989. Breeding soybean cultivars for increased seed protein percentage. In: World Soybean Research Conference IV. *Proceedings*, Buenos Aires. Actas, Buenos Aires: AASoja, 2, p. 1079-1085

Burton, J. W., C. Harlow, E. C. Theil. 1998. Evidence for reutilization of nodule iron in soybean seed development. *J. Plant Nutr.* 21 913–927 10.1080/01904169809365453

Caliskan, S., I. C. M. Ozkaya, M. Arslan. 2008. The effects of nitrogen and iron fertilization on growth, yield and fertilizer use efficiency of soybean in a Mediterranean-type soil. Field Crops Research 108:126-132. DOI:10.1016/j.fcr.2008.04.005.

Carlson, J.B. 1973. Morphology. In: Caldwell, B.E. (Ed.), Soybeans: Improvement, Production and Uses. Madison, Wisconson: USA.

Carpenter, J., A. Felsot, T. Goode, M. Hammig, D. Onstad, and S. Sankula. 2002. Comparative Environmental Impacts of Biotechnology-derived and Traditional Soybean, Corn, and Cotton Crops. Council for Agricultural Science and Technology, Ames, Iowa. www.cast-science.org. Sponsored by the United Soybean Board. www.unitedsoybean.org.

Carrera, C.S., C. M. Reynoso, G.J. Funes, M.J. Martinez, J. Dardanelli, S.L. Resnik. 2011. Amino acid composition of soybean seeds as affected by climatic variables. Pesq. agropec. bras., Brasília, v.46, n.12, p.1579-1587.

Carrera, C., M.J. Martínez, J. Dardanelli, M. Balzarini. 2009. Water deficit effect on the relationship between temperature during the seed fill period and soybean seed oil and protein concentrations. Crop Science, v.49, p.990-998.

Chapman, S.C., G. L. Hammer, D. W. Podlich, & M. Cooper. 2002. Linking bio-physical and genetic models to integrate physiology, molecular biology and plant breeding. p. 167-187 M.S.

Choung, M.G., I.Y. Baek, S.T. Kang, W.Y. Han, D.C. Shin and H.P. Moon .2001. Determination of protein and oil contents in soybean seed by near infrared reflectance spectroscopy. Korean Journal of Crop Science, 46 (2001), pp. 106–111

Clemente T.E. and. E. B. Cahoon .2009. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. Plant Physiology, November 2009, Vol. 151, pp. 1030–1040

Cregan, P.B. 1999. DNA Markers, Maps and Technologies. Proceedings of World SoybeanResearch Conference VI, pp. 46-58, Chicago, august 1999, Superior Printing, Champaign, Illinois, USA. Crossa, J. 1990. Statistical analysis of multi-location trials. Adv. Agro. 44, 55-86.

Crossa, J., Vargas, Eeuwijk, F.A., Jiang, C., Edmeades, G.O. and Hoisington, D. 1999. Interpreting genotype x environment in tropical maize using linked molecular markers and environm ental covariables. *Theoretical Applied Genetics* 99:611-625.

Crossa, J.; Cornelius, P.L. 1997. Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of errors variances. Crop Science 37: 406-415.

Crossa, J.; Gauch, H.G.; Zobel, R.W. 1990. Additive main effects and multiplicative analysis of two international maize cultivar trials. Crop Science 30: 493-500. CSO/MAL. 2009, 2010, 2011. Crop Forecast Survey. Lusaka: CSO

Dardanelli, J.L., M. Balzarini, M.J. Martinez, M. Cuniberti, S. Resnik, S.E. Ramunda, R. Herrero,H. Baigorri (2006): Soybean maturity groups, environments, and their interaction definemegaenvironments for seed composition in Argentina, Crop Sci 46: 1939–1947.

Dayaman, V. 2007. Diversity Analysis in Soybean [*Glycine Max* (L.) Merrill] using Morphological and Simple Sequence Repeat (SSR) Markers. Coimbatore: Department of Plant Molecular Biology and Biotechnology, Centre for Plant Molecular Biology.

Devi, K.N., L.N.K. Singh, M.S. Singh, S.B. Singh and K.K. Singh. 2012. Influence of Sulphur and Boron Fertilization on Yield, Quality, Nutrient Uptake and Economics of Soybean (*Glycine max*) under Upland Conditions. Journal of Agricutural Science. Vol.4, No.4.

Dordevic, V., J. Miladinovic, V. Dukic, M. Tatic, S. Balesevic-Tubic, G. Dozet, K.Petrovic.2010. Oil and protein content in soybean varieties. Proceedings of the 51th Oil industry Conference. Herceg Novi, June 27–July 02: 77-82.

Dornbos, D. L Jr. and R. E Mullen. 1992. Soybean Seed Protein and Oil Contents and Fatty-Acid Composition Adjustments by Drought and Temperature, *Journal of the American Oil Chemists' Society*, Vol. 69, pp. 228- 231.

Dugje, I.Y., L.O. Omoigui, F. Ekeleme, R. Bandyopadhyay, P. Lava Kumar, and A.Y. Kamara. 2009. Farmers' Guide to Soybean Production in Northern Nigeria. © International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria

Eberhart, S.A. and W.A. Russell, 1966. Stability parameters for comparing varieties. Crop Sci., 9: 36-40.

El-Shemy, H. A., Soybean and Nutrition. 2011. Published by InTech, Janeza Trdine 9, 51000 Rijeka, Croatia.

Evans, L. T. 1996. Crop Evolution, Adaptation and Yield. Cambridge Univ. Press.

Fabre, F. and C. Plancho. 2000. Nitrogen nutrition, yield and protein content in soybean. Plant Science 152 (2000) 51-58. Elsevier Science Ireland Ltd.

Falconer, D. S .1981. Introduction to quantitative genetics-2nd ed, Longman group limited, Burnt Mill, Harlow, Essex CM20 2JE, England.

FAO. 2014. Risk Assessment Report of the Genetic Modifiation Advisory Committee (GMAC) for an application for approval for release of product of A5547- 127 soybean by Bayer Co. (Malaysia) SDN. BHD.

FAO. 1999. Manual for training in seed technology and seed production. R. K. Triveds, I. A. Usman and J. C. Umeh (eds) FAO, Rome, 16p.

FAOSTAT. 2012; Available from: http://faostat.fao.org/. Accessed November, 2015.

Fasahat, P., A. Rajabi, S. B. Mahmoudi, M. A. Noghabi and J. M. Rad. 2015. An Overview on the Use of Stability Parameters in Plant Breeding, Biometrics & Biostatistics International Journal. Volume 2 Issue 5- 2015.

Fehr, W.R., J.A. Hoeck, S.L. Johnson, P.A. Murphy, J.D. Nott, G.I. Padilla, G.A. Welke. 2003. Genotype and environment influence on protein components of soybean. Crop Sci., 43:511-514.

Ferguson, R. B., C. A. Shapiro, A. R. Dobermann and C. S. Wortmann. 2006. Fertilizer Recommendation for soybeans. NebGuide G859. 6 pp.

Fikere, M., T. Tadesse and L.Tesfaye .2008. Genotype-Environment Interactions and Stability Parameters for Grain Yield of Faba Bean (*Vacia faba* L.) Genotypes Grown in South Eastern Ethiopia. Int. J. Sustain. Crop Prod. 3(6):80-87

Filho, M.M., D. Destro, L.M. Lilian Azevedo, W. A. Spinosa, Mercedes, C. Carrão-Panizzi and Montalvan R. 2001. Relationships among oil content, protein content and seed size in soybeans. Brazilian Archives of Biology and Technology, Vol 44, N.1: pp. 23-32.

Fox, P.N., J.Crossa, I. Ramagosa. 1997. Multi environment testing and genotypes x environment interaction. In: Kempton R.A., P.N. Fox, (Eds) Statistical methods for plant variety evaluation. London: Chapman & Hall, P.117- 138

Fukuda, Y. 1933. Cytogenetical studies on the wild and cultivated Manchurian soybeans (*Glycine* L.). Japanese Journal of Botany 6, 489–506.

Gallagher S. 2010. Modeling Protein Structure. Summer Research Program for Science Teachers. Harry S. Truman High School. Bronx.

Gauch G. 1992. Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs. Elsevier Science Publishers, Amsterdam, The Netherlands.

Gauch, H. G., and R.W. Zobel, 1996. AMMI analysis of yield trials. In: M. S. Kang and H. G.Gauch, Jr (Eds.). Genotype- environment-interaction. Boca Raton, CRC press. Pp 85-122.

Gauch, H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. Crop Science 46:1488-1500.

Gibson, L. and R. Mullen. 1996. Soybean seed composition under high day and night growth temperatures. JAOCS. 73: 733-737.

Gobarah M.E., M. H. Mohammad and M. M. Tawfik. 2006. "Effect of phosphorus fertilizer and foliar spraying with zinc on growth, yield and quality of groundnut under reclaimed sandy soils", Journal of Applied Science Research, Pakistan, vol. 2, pp. 491-496.

Gomez, K.A., and A.A. Gomez. 1984. Statistical procedures for agricultural research Second edition. John Wiley and Sons, Inc.: New York.

Grieshop, C.M. and G.C. Fahey. 2001. Comparison of quality characteristics of soybeans from Brazil, China and the United. Journal of Agricultural and Food Chemistr, Vol. 49, No.5, pp. 2669-2673.

Gurmu, F., H. Mohammed and G. Alemaw. 2009. Genotype x Environment Interactions and Stability of Soybean for Grain Yield and Nutrition Quality. African Crop Science Journal, Vol. 17, No. 2, pp 87-99.

Hartwig, E. 1973. Varietal development, in: Soybeans: Improvement, Production, and Uses, (Monograph 16), edited by B. E. Caldwell American Society of Agronomy, Madison, Wisconsin: pp. 187-210, Chap.6.

Havlin, L.J., Beaton, D.J., Tisdale, L.S. & Nelson, L.W. (1999). Soil fertility and fertilizers. *Prentice Hall of Indian*. 6th ed., pp: 220, 227, 228, 319-346.

Hayward, M.D., N.O. Bosemark, and I. Romagosa. 1993. Plant Breeding Principles and Prospects. Chapman and Hall: Great Britain.

Heidarzade, A., M. Esmaeili, R. Abbasi. 2016. Response of soybean (*Glycine max*) to molybdenum and Iron spray under well- watered and water deficit conditions. Journal of Experimental Biology and Agricultural Sciences. 4(1) 37 – 46.

Hellal, F.A. And M.T. Abdelhamid. 2013. Nutrient Management Practices for Enhancing Soybean (*Glycine max* L.) Production Acta Biologica Colombiana. 18 (2); 239-250.

Hollung, K, M. Overland, M. Hrustic, P. Sekulic, Miladinovic J, H. Martens, B. Narum, S. Sahlstrom, M. Sorensen, T. Storebakken, A. Skrede .2005. Evaluation of Non starch Polysaccharides and Oligosaccharide Content of Different Soybean Varieties (*Glycine max*) by Near-Infrared Spectroscopy and Proteomics. J. Agric. Food Chem. 53(23):9112-9121.

Hong, D.Y. and S. Blackmore (Eds). 2015. The Plants of China. Cambridge University Press.

Hymowitz T., Dudley J. W., Collins F. I., Brown C. M., 1974. Estimations of protein and oil concentration in corn, soybean, and oat seed by near-infrared light reflectance. Crop Sci. 14: 713–715.

Hymowitz, T. 1995. Evaluation of wild perennial Glycine species and crosses for resistance to Phakopsora. Urbana, IL: National Soybean Research Laboratory. pp. 33-37.

Hemantarajan, A. and S.K. Trivedi. 1997. Growth and yield of soybean (Glycine max L. Merr) as influenced by sulphur and iron nutrition. Indian J. Plant Physiol., 2: 304-306.

Hymowitz, T., and C.A. Newell. 1981. Taxonomy of the genus Glycine, domestication and uses of soybeans. Economic Bot. 35: 272-288.

IITA.2009. Soybean overview. Summary. 5pp

Isaza, M.I., J. Luquez, and M.E.Weilenmann. 2002. Variability of oil and protein contents and oil composition in soybean, sown in different environments of the southern part of Buenos Aires Province, in Argentina. INTA, Balcarce, Argentina.

Jamal, A., I. S. Fazli, S. Ahmad, M.Z. Abdin and S.J. Yun. 2005. Effect of Sulphur and Nitrogen Application on Growth Characteristics, Seed and Oil yields of Soybean Cultivars. Korean J.Crop Sci. 50(5): 340

Lance, G. and B. Garren. 2005. Origin, History, and Uses of Soybean (*Glycine max*). Iowa State University, Department of Agronomy.

Jandong, E.A., M.I. Uguru and B.C. Oyiga. 2011. "Determination of yield stability of soybean genotypes across diverse soil pH levels". Journal of Applied Biosciences, 2011, Vol 43: pp 2924 – 2941

Jaureguy, L. M., P. Chen, A. M. Scaboo. 2011. Heritability and Correlations Among Food-grade Traits in Soybean. Plant Breeding 130 (6): 647-652

Joshi, T., B. Valliyodan, J. H. Wu, S. H. Lee, D. Xu, H.T. Nguyen. 2013. Genomic differences between cultivated soybean, *G. max* and its wild relative *G. soja*. BMC Genomics 2013 :14 (Suppl 1):S5

Kabir R., S. Yeasmin, A. K. M. M. Islam and M. A. R. Sarkar. 2013. Effect of Phosphorus, Calcium and Boron on the Growth and Yield of Groundnut (*Arachis hypogea* L.) International Journal of Bio-Science and Bio-Technology Vol. 5, No. 3.

Kandel, H. 2015. Soybean Physiological Maturity. Crop and Pest Report; Plant Science. North Dakota State University. <u>https://www.ag.ndsu.edu/cpr/plant-science/soybean-physiological-maturity-09-10-15. Accessed 13.01.2017</u>.

Kobrae, S. and K. Shamsi. 2015. Relationships between Oil, Protein and Dry Matter in Soybean Seed with Some of Micronutrient Fertilization. Res. J. Soil Biol. ISSN 1819- 3498 / DOI: 10.3923/rjsb.2015

Kane, M.V., C.C. Steele, L.J. Grabau, C.T. Mackown, Hildebrand D.F.1997. Early-maturing soybean cropping system: III. Protein and oil contents and oil composition. Agron. J., 89:464-469.

Kang, M.S. and D.P. Gorman. 1989. Genotype X environment interaction in maize. Agronomy Journal 81: 662-664.

Koivisto, J. 2003. "Glycine Max." FAO: FAO Home. FAO Web. Accessed 25 May 2016.

Kolapo, A. 2011. Soybean: Africa's Potential Cinderella Food Crop, Soybean - Biochemistry, Chemistry and Physiology, Prof. Tzi-Bun Ng (Ed.), ISBN: 978-953-307-219-7, InTech.

Krishnan, H.B. 2005. Engineering soybean for enhanced sulfur amino acid content. Crop Science. 45: 454-461.

Kumar, V., A. Rani, and G.S. Chauhan. 2010. Nutritional Value of Soybean, in The soybean: botany, production and uses, G. Singh, Editor, CABI. p. 375-403.

Kumar, V., Rani, A., Solanki, S and Hussain, S.M. .2006. Influence of growing environment on the biochemical composition and physical characteristics of soybean seed. Journal of Food Composition and Analysis 19: 188-195.

Lee, J. H, J. H. Kwon, J. W. Kang, C. H. Park, B. S. Tae and S. W. Kim .2009. "Biodiesel Production from Various Oils under Supercritical Fluid Conditions by Candida antartica Lipase B Using a Stepwise Reaction Method," Applied Biochemistry and Biotechnology, Vol. 156, No. 1-3, , pp. 24-34.

Lee, J.D., J. G. Shannon and M. G. Choung .2011. Application of Nondestructive Measurement to Improve Soybean Quality by Near Infrared Reflectance Spectroscopy, Soybean – Applications and Technology, Prof. Tzi-Bun Ng (Ed.), ISBN: 978-953-307-207-4, InTech, Available from: http://www.intechopen.com/books/soybean-applications-and-technology/application-of-nondestructivemeasurement- to-improve-soybean-quality-by-near-infrared-reflectance-sp

Leffel, R.C., P.B. Cregan, A.P. Bolgiano, D.J. Thibeau. 1992. Nitrogen metabolism of normal and high-seed-protein soybean, Crop Sci. 32 (1992) 747–750.

L'Hocine, L. and J.I. Boye. 2007. Allergenicity of soybean: new developments in identification of allergenic proteins, cross-reactivities and hypoallergenization technologies. Critical reviews in food science and nutrition,. 47: 127-43.

Lin, C.S., M.R. Binns, L.P. Lefkovitch. 1986. Stability analysis: Where do we stand? Crop Sci., 26: 894–900.

Linnaeus, C., 1737, Linnean taxonomy. Sweden

Liu, K. 1997, Soybeans, Chemistry and Nutritional Value of Soybean Components, Springer U, Chapman and Hall. pp 25-113

Liu, K., F. Orthoefer, and E.A. Brown. 1995. Association of seed size with genotypic variation in the chemical constituents of soybeans. Journal of American Oil Chemistry Society. 72 (2), 189-192.

Lokuruka, M. 2010. Soybean Nutritional Properties: The good and the bad about Soy Foods Consumption-A Review. African Journal of Food Agriculture Nutrition and Development.Volume 10 No.4.

Lubungu, M., W. Burke and N.J. Nicholas. 2013. Challenges of Smallholder Soybean Production and Commercialization in Eastern Province of Zambia, Indaba Agricultural Policy Research Institute, Policy Brief.

Magagane, E. 2012. Genotype by Environment Interactions in Soybean for Agronomic Traits and Nodule Formation. MSc. Thesis. University of Limpopo, South Africa.

Mahama, O. 2011. Growth and Yield Response of early and medium maturity soybean (*Glycine max* (L) Merrill) varieties to row spacing, MSc. Thesis, Kwame Nkrumah Univrsity of Science and Technology, Kumasi, Ghana.

Maphosa, M.T. 2015. Growth, Nodulation and Yield response of promiscuous and nonpromiscuous soybean cultivars to inoculation in different soil types under glasshouse and field conditions. MSc. Dissertation, University of Limpopo.

Marjanovic-Jeromela, A., N. Nevena, G.V. Jelica, H. Nikola, K.S. Ankica, V. Mirjana and M. Radovan. 2011. Genotype by environment interaction for seed yield per plant in rapeseed using AMMI model, Pesq. agropec. bras. Brasília, v.46, n.2, p.174-181.

Marschner, H. 2011. Marschner's nutrition of higher plants. London: Academic Press.

Martin, F.W., 1988. Soybean: Echo Technical Note. ECHO, 17391 Durrance Rd., North Ft. Myers FL 33917, USA. Pp. 1-5.

Martin, F.W. 1984. Handbook of Tropical Food Crops. CRC Press: United States.

Masindeni, D.R. 2013. Evaluation of South African high quality protein maize (*Zea mays* L.) inbred lines under optimum and low nitrogen conditions and the identification of suitable donor parents. PhD Thesis, University of the Free State.

Miladinovic, J., Hrustic M., Vidic M. 2011. Soybean. Institute of Field and Vegetable Crops, Novi Sad and Sojaprotein, Bečej, AMB Graphics, Novi Sad. P. 510.

Miladinovic, J., H. Kurosaki, J.W. Burton, M. Hrustic, D. Miladinovic. 2006. The adaptability of short season soybean genotypes to varying longitudinal regions. Eur. J. Agron., 25: 243-249.

Miti, J. M. 1997. Soybean (*Glycine max* (L.) Merr.). In S.W. Muliokela (Ed.) Zambia seed technology handbook. Berlings, Arlov, Sweden.

Mpepereki, S. 2001. Soybean N2 fixation: A conceptual model for maximising benefits from research and extension for smallholder farmers. African Crop Science Conference Proceedings 5: 861-865.

Mpepereki, S., Javaheri, F., Davis, P. 2000. Soybean and Sustainable Agriculture: Promiscuous Soybean in Southern Africa. Field Crops Res 63, 137 – 149.

Mutsaers, H.J.W. 1991. Opportunities for second cropping in south eastern Nigeria. RCMD Research Monograph No 4, IITA, Ibadan, Nigeria, 28p.

Nahar, K. S., M. A. L. Ahmed, M. A. A. Akanda, Mondal and M. A. Islam. 2010. Genotype-Environment Interaction for Cob Yield and Maturity in Baby Corn (Zea mays L.), Bangladesh J. Agril. Res. 35(3): 489-496, September 2010

Ndambo, N. 2014. Investment Opportunity in Zambia's Agriculture Sector, Zambia National Farmers Union Publication.

Ngalamu, T., M. Ashraf and S. Meseka. 2013. Soybean (*Glycine max* L) Genotype and Environment Interaction Effect on Yield and Other Related Traits, *American Journal of* Experimental Agriculture, 3(4): 977-987, 2013.

Ngeze, P. B. 1993. Learn How to Grow Soybean. Nairobi, Kenya. CTA Publ. 21pp.

Ning, H., D. Zhang, S. Zhang, O. Yang .2003. Ecological Features of oil and protein content of soybean in Northeast China. Soybean Sci., 22: 132-136.

Nkhoma, N. 2013. Stability of Yield and Antioxidant Content of Selected Advanced Cowpea (*Vigna unguiculata* [L] Walp.) Mutation Derived Lines. MSc. Thesis, University of Zambia.

Noureldin, N.A, M.Z. Hassan, R.K. Hassaan, S.I. Abdel-Wahab .2002. Performance of some soybean in sandy soil as influenced by some abiotic stresses. II. Effect on seed yield and some yield attributes. Ann. Agric. Sci., 47: 209-223.

Nwokolo, E., and J. Smartt. 1996. Food and feed from legumes and oilseeds. Chapman and Hall.

Odewale, J. O., C. D. Ataga, C. Agho, G. Odiowaya, M. N. Okoye and E. C. Okolo. 2013. Genotype evaluation of coconut (*Cocos nucifera L.*) and mega environment investigation based on Additive Main effects and Multiplicative Interaction (AMMI) analysis, Research Journal of Agricultural and Environmental Management Vol. 2(1), pp. 001-010

OECD .2000. Consensus Document on the Biology of *Glycine Max* (L.) Merr. (Soybean), OECD Environmental Health and Safety Publications, Series on Harmonization of Regulatory Oversight in Biotechnology No. 15, Environment Directorate, Organisation for Economic Co-operation and Development, Paris.

O'Hara, G. W. 2001. Nutritional constraints on root nodule bacteria affecting symbiotic nitrogen

fixation: a review. Aust. J. Exp. Agric. 41 417–433 10.1071/ea00087

Ojo, D.K., M.A. Adebisi, B.O. Tijani, 2002. Influence of environment on protein and oil contents of soybeans seed (*Glycine Max* (L.) Merril). Global J. Agric. Sci., 1: 27-32.

Oliveira, B.A., J.B. Duarte and J.B. Pinheiro. 2003. Application of AMMI analysis in the assessment of yield stability in soybean. Brazilian Agricultural Research, Vol.38 no.3

Opperman, C. and N. Varia. 2011. Southern Africa Trade Hub, Technical Report: Soybean ValueChain,USAID/SouthernAfrica.<u>http://www.satradehub.org/agricultural-value</u> <u>chains/soybean-value-chain-report</u>

Osiru, M.O., O.M. Olanya, E. Adipala, R. Kapinga, B. Lemaga. 2009. Yield stability analysis of *Ipomoea batatus* L. cultivars in diverse environments. Aust J Crop Sci. 3(4):213-220.

Tukamuhabwa, P., M. Asiimwe, M. Nabasirye, P. Kabayi1 and M. Maphosa. 2012, Genotype by Environment Interaction of Advanced Generation Soybean Lines for Grain Yield In Uganda, African Crop Science Journal, Vol. 20, No. 2, Pp. 107 – 115.

Pacheco, R. M., J.B. Duarte, R. Vencovsky, J.B. Pinheiro, A.B. Oliveira .2005. Use of supplementary genotypes in AMMI analysis. Theor Appl Genet. 110: 812-818.

Padgette, S. R., N. B.Taylor, D. L.Nida, M. R. Bailey, J. MacDonald, L. R. Holden, R. L.Fuchs. 1996. The composition of glyphosate-tolerant soybean seeds is equivalent to that of conventional soybeans. *J. Nutr.*, *126*, 702-16.

Panthee, D.R., V.R. Pantalone, D.R. West, A.M. Saxton, and C.E. Sams. 2005. Quantitative trait loci for seed protein and oil contents, and seed size in soybean. Crop Sci. 45:2015–2022.

Pathan, S.M., T. Vuong, K. Clarke, J.D. Lee, J.G. Shannon, A. C. Roberts, M. R. Ellersieck, J. W. Burton, P. B. Cregan, D. L. Hyten, H. T. Nguyen, and D. A. Sleper. 2013. Genetic Mapping and Confirmation of Quantitative Trait Loci for Seed Protein and Oil Contents and Seed Weight in Soybean. Crop Sci. 53: 765-774.

Peters, J. W., Szilagyi R. K. 2006. Exploring new frontiers of nitrogenase structure and mechanism. *Curr. Opin. Chem. Biol.* 10 101–108 10.1016/j.cbpa.2006.02.019 [PubMed]

Peterson, C.J., P.S. Graybosch, P.S. Baenziger, A.W. Grombacher. 1992. Genotype and environment effects on quality characteristics of hard red winter wheat. Crop Sci 32(1): 98-103.

Pfeiffer, T.W., L.J. Grabau, J.H. Orf.1995. Early maturity soybean production system - genotype-X environment interaction between regions of adaptation. Crop Sci. 35(1):108-112.

Piepho, H.P. 1996. Analysis of genotype environment interaction and phenotypic stability. In: M.S. Kang, and H.G. Zobel Jr (Eds), pp. 151-174. Genotype by Environment Interaction. CRC Press, Boca Raton.

Piper, E.L., Boote K.J. 1999. Temperature and genotype effects on soybean seed oil and protein concentrations. J. Ame. Oil Chem. Society, 76:1233-1241.

Popovic, V., J. Miladinovic, M. Tatic, V. Djekic, G. Dozet, V. Dukic and N. Grahovac. 2013. Stability of soybean yield and quality components. African Journal of Agricultural Research. Vol.8 (45),pp. 5651-5658

Purchase, J. L., 1997. Parametric analysis to describe Genotype X Environment interaction and yield stability in winter wheat. Ph. D. Thesis, Department of Agronomy, Faculty of the Free State, Bloemfontein, South Africa.

Purchase, J.L., H. Hatting, and C.S. Van Deventer .2000. Genotype \times environment interaction of winter wheat in South Africa: II. Stability analysis of yield performance. S. Afr. J. Pl. Soil. 17: 101–107

Qiu, L.J. and R. Z. Chang. 2010. The Origin and History of Soybean, in The soybean Botany, Production and Uses. Singh G. Ed © CAB International. Oxfordshire, UK.

Ramana, M.V and A. Satyanarayana. 2006. Stability Analysis for Quality Characters in Soybean *Glycine Max* (L.) Merrill Legume Research, Vol. 29 (4): 274 – 277.

Rashidi, M., E. Farshadfar, M.M. Jowka. 2013. AMMI analysis of phenotypic stability in chickpea genotypes over stress and non-stress environments International Journal of Agriculture and Crop Sciences. Vol., 5 (3), 253-260.

Rienke, N. and N. Joke. 2005. Cultivation of soya and other legumes. Agrodok- series No. 10. Agromisa. CTA publication. 69 pages.

Rodrigues, J. I. S., Arruda, K. M. A., Cruz, C. D., Piovesan, N. D., Barros, E. G., and Moreira, M. A. 2014. Biometric analysis of protein and oil contents of soybean genotypes in different environments. *Pesq. Agropec. Bras.* 49, 475–482.

Rose, I.A. 1988. Effects of moisture stress on the oil and protein components of soybean seeds. Australian Journal of Agricultural Research, v.39, p.163-170.

Rotundo, J.L. and M.E. Westgate. 2009. Meta-analysis of environmental effects on soybean seed composition. Field Crops Research, v.110, p.147-156.

Sato, T., K. Eguchi, T. Hatano, and Y. Nishiba. 2008. Use of Near-Infrared Reflectance Spectroscopy for the Estimation of the Isoflavone Contents in Soybean Seeds, *Plant Production Science*, 11, 4, Oct 2008, 481-486, 1342-943X.

Sato, T.2011. Estimation of Fatty Acid Composition in Soybean Powder by Examining Near Infrared Spectroscopic Patterns, Soybean - Biochemistry, Chemistry and Physiology, Prof. Tzi-Bun Ng (Ed.), ISBN: 978-953-307-219-7, InTech, Available from: <u>http://www.intechopen.com/books/soybean-biochemistrychemistry-and</u> physiology/estimationof-fatty-acid-composition-in-soybean-powder-by-examining nearinfrared- spectroscopic-patt Schoeman, L.J. 2003. Genotype x environment interaction in sunflower (Helianthus annuus) in South Africa. M.Sc. Thesis, Department of Plant Sciences (Plant Breeding), Faculty of Natural and Agricultural Sciences, University of the Free State, South Africa.

Schwender J, J.B. Ohlrogge, Y. Shachar-Hill .2003. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing Brassicanapus embryos. J Biol Chem 278: 29442–29453.

Sharma B.M. and J. S. P. Yadav. 1997. Availability of phosphorus to grain as influenced by phosphatic fertilization and irrigation regimes. Indian Journal of Agricultural Science, vol. 46, pp. 205-210.

Shurtleff, W. and A. Aoyagi .2004. History of Soy Flour, Grits, Flakes, and Cereal-Soy Blends -Part 1.Soyinfo Center. In the Unpublished Manuscript, *History of Soybeans and Soyfoods*, *1100 B.C. to the 1980s*. Soyfoods Center, Lafayette, California.

Silva, K.B., A.T. Bruzil, A.M. Zuffo, E.V. Zambiazzi, I.O. Soares, P.M. de Rezende, V. Fronza, G.D.L. Vilela1, F.B.S. Botelho, C.M. Teixeira and M.A. de O. Coelho. 2016. Adaptability and phenotypic stability of soybean cultivars for grain yield and oil content. Genet. Mol. Res. 15 (2): gmr. 15026756

Singh, G. (2010). The soybean: botany, production and uses: The origin and history of soybean, L. J. Qui & R. Z. Chang, 1-23, CAB International.

Singh, O. and Chaudhry, C.D. 1985b Stability analysis for protein and oil content in soybean. *J. Oilseeds Research.* **2:** 57-61.

Skvortzow, B.V. 1927. The soybean-wild and cultivated in Eastern Asia. Proc. Manchurian Res. Soc. Publ. Ser.A. Natual History. History Sect. No. 22:1-8.

Soyatech. 2012. Soy facts. [cited 20-01-2014]; Available from: http://www.soyatech.com/soy_facts.htm.

Soystats, 2013; A publication of the soybean association of America

Steele, R.D.G., J.H. Torrie, and D.A. Dickey. 1997. Principles and Procedures of Statistics: A Biometrical Approach. 3rd ed. McGraw-hill: New York.

Sudaric, A., M. Vrataric, D. S. Mladenovic, and Z. Zdunic. 2011. Genetic Improvement: Molecular-Based Strategies, In: Sudaric, A. (Ed) Soybean Molecular Aspects of Breeding. Published by InTech, Croatia.pp57. Doi: <u>www.intechopen.com</u>

Sun, Y.K. 2011. Identification of Soybean Seed Oil QTLs with Little or No Impact on Seed Protein. MSc. Thesis, University of Nebraska.

Suri, S. 2006. Bioinformatics. APH Publishing Corporation. New Dlehi- 110 002

Swanby Haidee, 2010, The Gates Foundation and Cargill push Soya onto Africa, African Center for Biosafety (ACB), Briefing Paper No. 20, South Africa. <u>www.biosafetyafrica.net</u>

TechnoServe. 2011. Southern Africa Regional Soybean Roadmap – Zambia Value Chain Analysis, <u>www.technoserve.org</u>. Accessed March 17, 2015.

Tefera, H. 2011. Breeding for Promiscuous Soybeans at IITA, Soybean - Molecular Aspects of Breeding, Dr. Aleksandra Sudaric (Ed.), ISBN: 978-953-307-240-1, InTech, Available from: http://www.intechopen.com/books/soybean-molecular-aspects-of-breeding/breeding-forpromiscuous-soybeans-at-iita cited January, 12 2015.

Thoenes, P. 2006. Background paper for the Competitive Commercial Agriculture in Sub-Saharan Africa (CCAA) Study. Soybean: International Commodity Profile. Food and Agriculture Organization of the United Nations.

Thorne, J., W.R. Fehr.1970. Incorporation of high-protein, exotic germplasm into soybean populations by 2- and 3- way crosses. Crop Sci. 10, 652- 655.

Truberg, M. and H.G. Huhn. 2000. Contribution to the analysis of Genotype Environment Interactions: Comparison of Different parametric and Non-parametric test for interactions with emphasis on crossover interactions. J. Agro. Crop Sci. 185,267-274

Tubic, B. S., V. Dordevic, J. Miladinovic, V. Dukic, J., Miladinovic, V. Dukic, and M. Tatic. 2011. Stability of soybean seed composition. Genetika, Vol 43, No.2, 217 -227.

Tukamuhabwa, P., M. Asiimwe, M. Nabasirye, P. Kabayi and M. Maphosa .2012. Genotype by Environment Interaction of Advanced Generation Soybean Lines for Grain Yield in Uganda, African Crop Science Journal, Vol. 20, No. 2, Pp. 107 – 115.

Tyagi, S.D., M.H. Khan .2010. Genotype x environment interaction and stability analysis for yield and its components in Soybean ((Glycine max L.) Merrill). Soybeans genetics newsletter v.37

Uncu, A.H, H.H. Arioglu. 2005. The effect of some growth regulators on seed yield and quality characteristics of soybean at different planting times in second crop soybean farming, The Sixth Field Crops Congress of Turkey. Proc., 1: 375-380.

<u>USDA National Nutrient Database for Standard Reference, Release 21"</u> (PDF). United States Department of Agriculture. Accessed 2017-01-13.

Ustun, A., F.L., Allen, B.C. English .2001. Genetic progress in soybean of the U.S. Midsouth, Crop Science, 41: 993–998.

Van Schoote, M. 2011. Breeding of food grade soybeans, MSc Thesis, Ghent University.

Vollmann, J., C.N. Fritz, H.Wagentrist and P. Ruckenbauer. 2000. Environmental and Genetic Variation of Soybean Seed Protein Content under Central European Growing Conditions. Journal of the Science of Food and Agriculture 80: 1300-1306.

Wehrmann, V.K., W.R. Fehr, S.R. Cianzio, J.F.Cavins. 1987. Transfer of high seed protein to high-yielding soybean cultivars. Crop Sci. 27, 927-931.

Win M., S. Nakasathien and E. Sarobol. 2010. Effects of Phosphorus on Seed Oil and Protein Contents and Phosphorus Use Efficiency in Some Soybean Varieties. Kasetsart J. (Nat. Sci.) 44: 1-9.

Win M., S. Nakasathien and E. Sarobol. 2010. Effects of Phosphorus on Seed Oil and Protein Contents and Phosphorus Use Efficiency in Some Soybean Varieties. Kasetsart J. (Nat. Sci.) 44: 1-9.

Wilcox, J.R. 1985. Breeding soybean for improved protein quantity and quality. In: World Soybean Research Conference, III: Proceedings. Ed. R Shibles, Boulder, CO: Westview Press, p. 380-386.

Wilcox, J.R., J.F. Cavins. 1995. Backcrossing high seed protein to a soybean cultivar. Crop Sci. 35 1036-1041.

Wolf, R.B., J.F. Cavins, R. Kleiman, and L.T. Black. 1982. Effect of Temperature on Soybean Seed Constituents: Oil, Protein, Moisture, Fatty Acids, Amino Acids and Sugars. Journal of American Oil Chemists' Society 59:230-232.

Yaklich, R. W., B. Vinyard, M. Camp, S. Douglass. 2002. Analysis of seed protein and oil from soybean Northern and Southern Region Uniform Tests. Crop Science, *42*, 1504-1515.

Yamagata, M., Kouchi, H., Yoneyama, T., 1987. Partitioning and utilization of photosynthate

produced at different growth-stages after anthesis in Soybean (Glycine max-L Merr) - Analysis by long-term C-13 labeling Experiments. J. Exp. Bot. 38, 1247-1259.

Yan, W., M. S. Kang, B. Ma, S. Woods, and P. L. Cornelius. 2007. GGE Biplot vs. AMMI Analysis of Genotype-by-Environment Data. Crop Sci 47: 643-653

Yan, W. and I. Rajcan. 2002. Biplot evaluation of test sites and trait relations of soybean in Ontario. Crop Sci. 42(1): 11-20.

Zhang, F, D.H. Lynch and D.L. Smith.1995. Impact of low root temperatures in soybean [Glycine max (L.) Merr.] on nodulation and nitrogen fixation. Environ Exp Bot 35:279-285

Zhang, F., D.L. Smith. 1999. Soybean [*Glycine Max* (L.) Merr.], Physiology and Symbiotic Dinitrogen Fixation, Crop Yield, in Physiology and Processes, pp 375- 399. © Springer- Verlag Berlin Heidelberg.

Zhang, M.Q., M.S. Kang, P.F. Reese Jr, H.L. Bhardwaj .2005. Soybean cultivar evaluation via GGE biplot analysis. J. New Seeds 7: 37-50.

Zilic, S., V. Hadzitaskovic Sukalovic, M., Srebric, D. Dodig, M. Maksimovic, S. Mladenovic Drinic and M.Crevar .2009. Chemical compositions as quality parameters of ZP soybean and wheat genotypes. Genetika *41* (3): 297-308.

Zobel, R.W., J.W. Madison and H.G. Gauch Jr .1988. Statistical analysis of a yield trial. Agronomy Journal 80: 388-393.

CHAPTER 8 APPENDICES

APPENDIX 1: Correlations among soybean oil content and environmental parameters

						Org.										
	oil%	Rainfall	Temp	pН	Ν	Matter	Р	K	Na	Ca	Mg	Cu	Fe	Mn	Zn	S
oil%	1.000															
Rainfall	0.495	1.000														
Temp	0.383	-0.231	1.000													
рН	-0.368	-0.763	0.388	1.000												
Ν	0.131	0.843	-0.493	-0.921	1.000											
Org. Matter	0.161	0.019	0.859	0.142	-0.110	1.000										
Р	0.274	0.313	-0.666	-0.229	0.181	-0.864	1.000									
K	0.060	-0.015	0.515	0.641	-0.444	0.455	-0.188	1.000								
Na	0.139	0.264	0.316	0.419	-0.185	0.325	0.009	0.950	1.000							
Ca	0.164	-0.143	0.757	0.667	-0.560	0.643	-0.383	0.949	0.827	1.000						
Mg	-0.272	-0.283	0.482	0.794	-0.559	0.473	-0.384	0.932	0.813	0.895	1.000					
Cu	-0.829	-0.487	0.048	0.690	-0.395	0.209	-0.451	0.495	0.370	0.413	0.761	1.000				
Fe	0.839	0.866	0.032	-0.760	0.644	0.077	0.318	-0.130	0.080	-0.128	-0.456	-0.819	1.000			
Mn	0.241	-0.544	0.924	0.647	-0.783	0.641	-0.542	0.501	0.243	0.737	0.527	0.154	-0.241	1.000		
Zn	-0.485	-0.526	0.178	0.920	-0.704	0.069	-0.122	0.757	0.627	0.669	0.890	0.822	-0.709	0.382	1.000	
S	-0.295	0.613	-0.771	-0.698	0.884	-0.357	0.257	-0.436	-0.191	-0.641	-0.431	-0.046	0.258	-0.943	-0.410	1

	Protein	Dain fall	Tomp	ъН	N	Org. Mottor	D	K	No	Ca	Ma	Cu	Fo	Mn	Zn	S
Protein %	1 000	Kalli lali	Temp	pn	1	Watter	1	К	114	Ca	Ivig	Cu	rt	WIII	211	6
Poin foll	-0.928	1 000														
Kain tan Tomp	0.205	0.231	1.000													
1 emp	0.205	-0.231	1.000													
рН	0.810	-0.763	0.388	1.000												
Ν	-0.748	0.843	-0.493	-0.921	1.000											
Org. Matter	0.115	0.019	0.859	0.142	-0.110	1.000										
Р	-0.473	0.313	-0.666	-0.229	0.181	-0.864	1.000									
K	0.183	-0.015	0.515	0.641	-0.444	0.455	-0.188	1.000								
Na	-0.077	0.264	0.316	0.419	-0.185	0.325	0.009	0.950	1.000							
Cu	0.255	-0.143	0.757	0.667	-0.560	0.643	-0.383	0.949	0.827	1.000						
Mg	0.500	-0.283	0.482	0.794	-0.559	0.473	-0.384	0.932	0.813	0.895	1.000					
Ca	0.770	-0.487	0.048	0.690	-0.395	0.209	-0.451	0.495	0.370	0.413	0.761	1.000				
Fe	-0.969	0.866	0.032	-0.760	0.644	0.077	0.318	-0.130	0.080	-0.128	-0.456	-0.819	1.000			
Mn	0.460	-0.544	0.924	0.647	-0.783	0.641	-0.542	0.501	0.243	0.737	0.527	0.154	-0.241	1.000		
Zn	0.686	-0.526	0.178	0.920	-0.704	0.069	-0.122	0.757	0.627	0.669	0.890	0.822	-0.709	0.382	1.000	
S	-0.446	0.613	-0.771	-0.698	0.884	-0.357	0.257	-0.436	-0.191	-0.641	-0.431	-0.046	0.258	-0.943	-0.410	1.000

APPENDIX 2: Correlations among soybean protein content and environmental parameters
APPENDIX 3: Stepwise Regression: Oil versus Rainfall, Temp and Edaphic factors

Alpha-to-Enter: 0.25 Alpha-to-Remove: 0.3

Response is Oil on 14 predictors, with N = 75

Step	1	2
Constant	15.38	15.45
Fe	0.370	0.531
T-Value	7.93	5.80
P-Value	0.000	0.000
Rainfall		-0.00145
T-Value		-2.03
P-Value		0.046
Summary	0.981	0.961
R-Sq	46.29	49.20
R-Sq(adj)	45.56	47.78

APPENDIX 4: Stepwise Regression: Protein versus Rainfall, Temperature and Edaphic factors

Alpha-to-Enter: 0.3 Alpha-to-Remove: 0.4

Response is Protein on 14 predictors, with N = 75

Step	1	2
Constant	39.87	42.23
Р	-0.142	-0.055
T-Value	-2.46	-1.18
P-Value	0.016	0.244
S	-0.157	-0.082
T-Value	-2.23	-1.47
P-Value	0.029	0.145
Fe	-0.609	1
T-Value	-6.94	
P-Value	0.000	
Summary	2.21	1.72
R-Sq	17.04	50.57

	17101	00107
R-Sq(adj)	14.74	48.48