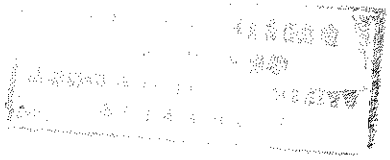
MULTIVARIATE ASSESSMENT OF ENVIRONMENTAL EFFECTS ON GRAIN YIELD AND COMPONENT CHARACTERS OF TEF (*Eragrostis tef* (*zucc.*) *Trotter*) GENOTYPES

BY

LEGESSE KASSA



A Thesis Submitted in (Part) Fulfillment for the Degree of Master of Science in Statistics in the Addis Ababa University

Addis Ababa, Ethiopia, June, 1997

TO THE MEMORY OF MY DAD

KASSA DEBUSHO



BLESSED BE GOD, WHO HAS NOT TURNED AWAY MY PRAYER, NOR HIS MERCY FROM ME !

PSALMS 66:20

ACKNOWLEDGMENT

I am very grateful to Dr. Desta Hamito, my research main advisor for his invaluable comments and suggestions which contributed to the successful realization of the study. My deep appreciation and thanks also goes to my co-advisor and instructor, Dr. Tadewos Koroto, for his critical reviewing of my work and making necessary comments and suggestions.

I am highly indebted to the department of statistics for the provision of computer facilities during off hours which has expedited the completion of the preparation of the thesis manuscript. I would like to appreciate the sincere effort made by members of the department's graduate council (DGC) to fine-tune the research proposal.

I would like to express my gratitude to members of Tef Improvement Programme and the Soil Section of Debre Zeit Agricultural Research Center, especially to Dr. Hailu Tefera, leader of Tef Improvement Programme and Selamyihun Kidanu, for their comments in the preparation of the thesis manuscript and provision of summarized data and useful reference materials. My thanks also go to National Meteorological Services Agency for producing summarized climate data.

I am also grateful to Van Lill, D. and Ms Marie Smith, from Rebulic of South Africa, for provision of reference materials through e-mail and by post from the beginning to the end of my thesis research.

I am particularly grateful to Ato Solomon Cheffo and , Solomon W/Mariam staff members in the Statistics Department, without whom the whole process of typing the thesis would have not been possible. My deep appreciation is also extended to Solomon Atnafu, staff member in the Mathematics

Department of A.A.U., who helped me in typing and printing the proposal. I am also indebted to Sisay Feseha, who facilitated and carried out part of the statistical analyses of the data using the computer.

I am very much indebted to my family members, especially to my brother Tadesse Kassa, without whose encouragement my work would not have reached this stage. I am pleased to take this opportunity to appreciate w | ro Abebech Abay for her deep concern and motherly treatment throughout my second year of study.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENT i

ABSTRACT vii

CHAPTER ONE 1

INTRODUCTION 1

CHAPTER TWO 6

REVIEW OF RELATED LITERATURE 6

General Literature Review 6

Literature specific to tef 11

CHAPTER THREE 14

METHODS OF STUDY 14

3.1. The Study Areas and Sources of Data 14

3.2. Variables Included in the Study 17

3.3 Statistical Procedures Used 19

3.3.1. Analysis of Variance 20

3.3.2. Canonical Correlation Analysis 24

3.3.3. Principal Component Analysis 29

Page

CHAPTER FOUR 34

RESULTS AND DISCUSSION 34

4.3. Principal Component Analysis 44

4.1. Analysis of Variance 34

4.2. Canonical Correlation Analysis (CCA) 37

CHAPTER FIVE 52

CONCLUSIONS AND RECOMMENDATIONS 52

REFERENCES 55

APPENDIX 60

List of Tables

Table	Page
3.1. NYT codes and Genotype codes of tef used in the study	16
3.2. Descriptions of Variables	18
3.3. Outline of the Combined Analysis of Variance over e environments based on RCB design with g genotypes	23
4.1. Analysis of variance table for grain yield of tef	35
4.2 Variance components and their ratios estimated for environment and genotype main effects and their interactions for grain yield and component characters	37
4.3. Canonical correlation coefficients between grain yield and component characters of 12 genotypes of tef and a set of environment variables	36
4.4. PCs analysis of yield and component characters data among genotypes	51
A.1. Normality test of grain yield and component characters of 12 tef genotypes	60
A.2. Combined analysis of variance for grain yield and component characters in 12 tef genotypes across 12 environments and 4 replications	63
A.3. Correlation coefficients among environmental variables measured in 1991/92, 1992/93 and 1993/94 crop seasons across four locations	64
A.4. Chi-square tests of the joint nullity of the smallest 5 - k canonical correlations	65
A.5. Contribution of individual canonical correlation	67

Table Page

A.6. Canonical variate analysis of grain yield and component characters for 12 tef genotypes and environment variables 69

A.7. Principal components analysis of yield and component characters for 12 tef genotypes grown under 12 environments 77

ABSTRACT

This study was designed to attain a better understanding of environmental factors such as soil and climate that could account for variation in grain yield and component characters of tef genotypes. Furthermore, it was intended to determine the source of variability in the relationships among grain yield and component characters within and among tef genotypes. In the study, 7 and 4 variables of soil and climate, respectively were collected from four locations for the crop seasons 1991/92, 1992/93 and 1993/94. Five plant variables of 12 tef genotypes were also collected for the same locations and crop seasons. Combined analysis of variance, analysis of variance components, canonical correlation analysis and principal component analysis were used to analyse the data.

The results of the study, based on combined analysis of variance and analysis of variance components, revealed that soil and climate variables accounted for a large portion of the variation in grain yield and component characters of tef genotypes. The results of canonical correlation analysis also supported this fact. The results of principal component analysis showed that each of the genotype variables considered in this study had their own share of contribution to the variation within and among genotypes.

CHAPTER ONE

INTRODUCTION

Tef, (*Eragrostis tef* (Zucc.) Trotter) is an indigenous crop to Ethiopia, and is by far the most important cereal crop. It is considered as the national staple food crop of the country. According to Vavilov (1951) tef originated in Ethiopia and became known to the Europeans through Portuguese contact in the 16th century (Taddesse, 1975). However, in some countries like South Africa, Kenya, Uganda, India, and the United States of America, tef is grown principally as animal feed (Tareke, 1972).

Tef is an important rain - fed crop and adapted to a wide range of climatic conditions and soil types. It is predominantly cultivated on sandy loam to black soil (Taddesse , 1969).Tef can be grown from sea level up to 2800m, but performs better at an altitude range of 1800 - 2100 meters above sea level, annual rainfall of 750 - 850 mm, growing season rainfall of 450 - 550 mm, and a temperature range of 10^0 C - 27^0 C (Seyfu, 1993b).

Tef has several uses and economic importance which make it very popular in the country. It is predominantly grown as a grain for human consumption. The most popular type of Ethiopian flat bread " Injera "is made from tef grain flour. Tef is also used for making " porridge ", and beverages like the local beer (Tella) and a local alcoholic beverage called " araki " or " katikala " (Taddesse, 1969).

Not only the tef grain but also its straw which is locally known as " chid " has valuable economic uses. Tef straw is used as feed for cattle especially during the dry season when there is acute shortages of feed resources. Tef straw is also used for construction purpose, that is, reinforcing mud for plastering walls of houses and other household items.

Nutritionally , tef plays a great role in Ethiopian diet. It is considered as the most nutritious cereal grain crop grown in Ethiopia (Rouk and Mengesha, 1963). It is a good source of vitamins and minerals, especially iron and calcium (Mengesha, 1966). Furthermore, it is believed that tef provides as much as two - thirds of the protein intake of the daily Ethiopian diet (Seyfu, 1988).

Due to the aforementioned versatile features, the crop is grown extensively throughout the country.

According to the report from the Central Statistical Authority (1996), the total land under the major crops, cereals, pluses and other crops , for the 1995/96 crop seseaon was 7.95 million hectares. Of these, 6.65 million is allocated for cereals. Tef, the single dominant cereal, occupies 2.10 million hectares constituting about 26% of the area under major crops and about 31% that of cereals.

According to Hofstra (1972) growth in plants is generally the result of interaction of many complex processes each of which is influenced either directly or indirectly by environmental factors. He noted that plants depend for their growth and development firstly on their genetic constitution and secondly, on their environment which is in the broadest sense includes soil and climatic factors. It is

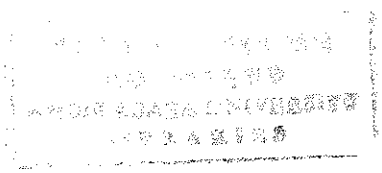
widely recognized that improvement in plant types can make a very significant contribution to higher grain yields.

Grain yield, an extremely complex character is dependent on a number of component characters. Moreover, these characters and yield itself are governed by a large number of factors including environmental causes (Katiyar, *et al*, 1979). Hence, studies are necessary to ascertain relationships among plant characters (traits) of tef genotypes, to evaluate the contribution of the characters that interact and affect yield and to determine the magnitude of influence of environmental factors on grain yield and traits.

Scientific research on tef started in the late 1950's by collecting land races from different parts of Ethiopia (Tareke (1974) as quoted by Hailu, 1988). So far, not much work has been done to generate information on character associations and to determine the relative importance of characters under different environments. Therefore, there is a need to generate knowledge about character relationships in tef and their relative merits so as to utilize them as a tool in the formulation of selection criteria in the improvement of tef.

The Debre Zeit Agricultural Experiment Station (now the Debre Zeit Agriculture Research Center (DZARC)), which is the coordinating center for tef research in Ethiopia, has been conducting trials on a number of parameters on tef since 1957 and has released a number of varieties to date.

As quoted by Hailu, 1988 selection was believed to be the best tool at hand for yield improvement in tef (Mulugeta, 1973). This belief arose from the earlier assumption that tef would only self-fertilize



and breed true. But this was found not be true. Tareke Berhe (1972) was able to cross fertilize tef and he thus has discovered an additional technique for improving tef varieties. The technique has not yet been applied in a wide scale.

The DZARC has, therefore, been engaged in improving tef varieties by selection and hybridization that would give better yield than local varieties traditionally used by farmers. The research trials have enabled farmers of Ethiopia to receive varieties (cultivars) with better yields. The DZARC research reports (1955 - 1994) shows that there are yield differences among genotypes over different sites. These differences of yields were presumed to have been brought about by environmental factors. However, the effects of the different environmental factors on yield have not been clearly determined.

The statistical analyses used so far in tef research as related to various environmental factors were univariate analyses (analysis of one response variate at a time). In examining environmental effects on yield and its component characters, cereal researchers measure yield and its component characters and attempt to develop relationships with associated environmental variables. As with many types of biological data, these characters respond concurrently to environmental effects and often correlated. Therefore, efforts to identify a single variable representative of a complex set of characters by using univariate analysis may be inadequate and has often proved to provide redundant information (Lezzoni & Pritts, 1991). Thus, the aim of this project is to study the effects of environmental factors on grain yield and component characters and the relationships among grain

yield and component characters to explain the variation observed in tef genotypes using multivariate statistical techniques. The main objectives of this study are :

1. To determine the relative contribution of genotype, environment and genotype- by-environment interaction, to the variation in grain yield and component characters of 12 genotypes tested in 12 environments;
2. To study the effects of a set of environmental factors such as soil and climate on variation in grain yield and component characters in tef genotypes, and
3. To determine the source of variability in the relationships among grain yield and component characters under different environmental conditions and genotypes.

The plan adopted is the following. Chapter two deals with the related literature. That is, it summarizes the statistical techniques used in related studies and some findings with regard to tef. The third chapter discusses methods of study. This part deals with the source of data, variables included and statistical techniques used. The study results and their discussions are given in chapter four, and chapter five contains the conclusions of the study.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

General Literature Review

It is a common practice in genetical studies and in trials of varieties and breeding lines to grow a series of genotypes in a range of different environments. If the genotypes all respond similarly to all the environments tested, their relative performance in other environments may be predicted with some confidence. More often, however, there are differences in the relative performance and an analysis of the pattern of the variation both between and within the genotypes is necessary to determine which particular genotypes respond differently from the others (Freeman & Dowker, 1973).

Various statistical techniques have been used to assist in the analysis of variation between and within genotypes. The first stage has usually been to regard the data as coming from a two-way table, with genotypes and environments being considered as rows and columns; a straightforward analysis of variance is then performed, taking out main effects of genotypes and environments and the genotype-environment interaction. It has been further observed by various authors as quoted by Freeman and Dowker, 1973, e.g. Yates and Cochran (1938), Finlay and Wilkinson (1963), Rowe and Andrew (1964), Eberhart and Russell (1966), Perkins and Jinks (1968), Breese (1969), Baker (1968), that the relation between the performance of different genotypes in the various environments and some measure of these environments is frequently linear, or nearly so. The technique, known as *joint regression analysis*, consists of regressing genotypes effects on environmental means and

examining the difference between genotypes both in their linear regression coefficients and deviations from linearity. This method has been criticized on statistical grounds by Freeman and Perkins (1971). The reason was that the fundamental statistical assumptions have not been satisfied in any of the work quoted. Furthermore, the lines that have been drawn account for the observed results, but it is not valid to regard them as regression lines or to compare their slopes statistically.

An apparently completely different approach to the analysis of variation between and within genotypes is by means of multivariate analysis (Freeman & Dowker, 1973). Multivariate techniques have usually been applied to groups of individual on which a whole set of measurements have been taken. All the measurements of a set are usually made on one individual of the group, but there is no theoretical need for this, and it is perfectly legitimate to regard the set of measurements on different individuals of a genotype. The data can then be examined by any method of multivariate analyses, one possible method being *principal component analysis.* In this method, components of variation within genotypes are taken out successively so that each extracts the maximum possible variation remaining in the system and all are orthogonal to each other.

It is better if cereal researchers measure complex traits, such as grain yield, plant height, panicle length, days to heading, etc. and develop relationships among the traits and environmental variables. Each complex trait, however, is a composite of individual traits that often vary due to different environmental factors. Identifying a single variable representative of the complex set of traits may not be possible, so the researcher is faced with the possibility of separately examining many related variables. If the researcher uses univariate statistics to quantify differences or

number of separate analysis required will be equal to the number of variables measured. With many types of biological data, however, correlation (multicollinearity) among variables is common, and information provided by separate univariate analyses will be redundant (Lezzoni & Pritts, 1991). Principal Component (PC) analysis is a method that can be used to identify patterns in a data set and to eliminate redundancy in univariate analyses when multicollinear data are involved.

The classical method of summarizing the variation in response of a number of varieties grown in different environments by partitioning the sum of squares into components due to varieties, environments and variety - environment interaction conveys little information on the individual patterns of response (Byth *et al.*, 1976). Further partition of the interaction sum of squares by grouping varieties or environments may help to identify major sources of interaction in the data, but when such groupings are achieved they are often difficult to interpret biologically (Silvey, 1982, as quoted by Kempton, 1984).

Williams *et al.* (1977) used PC analysis to define three components of growth and yield in sorghum and to correlate these characters with environmental variables. The three components were defined as crop density, total photosynthate, and vegetative extension. They found that the crop density component was significantly correlated with rainfall in three cultivars, and the total photosynthate components was significantly correlated with maximum temperature.

Preston *et al.* (1992) used PC analysis to determine the interrelationships of the quality attributes from a collection of 33 wheat samples grown in 14 countries. Van Lill *et al.* (1995) also used PC

analysis to determine the source of variability among yield and quality attributes over 16 environments and five cultivars.

Jenning (1979) measured various yield components in raspberry to investigate a relationship between cane diameter and the development of multiple lateral shoots and fruit nodes on these canes using PC analysis. The PC analysis identified two vectors of which the largest, termed the vigor vector , was associated with high expression of cane diameter and other vegetative characters. The second vector, called the compensation vector, described the tendency of high cane diameter and large number of multiple laterals to be compensated for by reduced development of fruiting nodes.

Yet another use of PC analysis has been to relate ecotypes and species habitats to specific environmental variables (Nevo *et al.*, 1979). Jensen and Hancock (1982) evaluated three species of strawberry from several populations collected from a diversity of environments to examine relationships among populations and communities. They found that combinations of morphological traits that are associated with the various community environments have evolved.

Pritts and Hanock (1984) (as quoted by Lezzoni and Pritts, 1991) used PC analysis to consolidate environmental and edaphic variables from sites that supported natural populations of lowbush blueberries. A second PC analysis was also performed to reduce the dimensionality of a group of growth and reproductive attributes that were characteristic of the blueberry populations. The PCs from the environmental variables were then correlated with the PCs from the growth and

reproductive variables, resulting in four significant correlations. These correlations among 10 PCs were much easier to interpret than 210 separate correlation coefficients.

Interpretation of a canonical correlation study can be formalized, and hence simpler than that for conventional correlation analysis. in conventional correlation, If for instance there are two sets of variables p and q, in conventional correlation , there will be p(p-1)/2 interrelations of the first set and q(q-1)/2 interrelations of the second set. In addition, there will be pq interrelations of the two sets of variables. In canonical correlation there are only q correlations ($q \leq p$). The latter ones, have the additional advantage of being independent of one another. Also variables within any canonical vector are independent. Therefore, interpretation is more direct (Pimentel, 1979).

Van Lill *et al.* (1995) used canonical correlation analysis to examine the relationship between variation in the biochemical composition of wheat flour and bread making quality. In addition, they used canonical variate analysis to determine the impact of groups of environmental variables such as climatic and edaphic factors on variation in yield, biochemical and rheological characteristics. Graybosch *et al.* (1993) used canonical correlation analysis to ascertain the extent to which a set of biochemical measurements was related to a set of quality measurements of wheat and to determine the particular components that have been responsible for these correlations. For discrimination between quality types in breeding material , where differences between groups are of more significance than differences between individual breeding lines of wheat, canonical variate analysis was employed (Osborne *et al.*, 1993, as quoted by Van Lill *et al.*, 1995).

Literature specific to tef

In 1951, Vavilov, recognized that the species was among a number of cultivated plants found only in Ethiopia. Collection and identification of cultivars was done by Melake-Hail (1964) and Tadesse (1975). Studies of the development of embryo sac and embryo of tef was carried out by Melake-Hail (1966). Tadesse (1969) dealt with the uses of tef in the Ethiopia society. Although as quoted by Tadesse, 1975, Clayton *et al.* (1974) had described the morphology of tef, a more detailed botanical characterization and morphological description of tef was made by Tadesse (1975). Tareke (1974) dealt with the breeding techniques, and as quoted by Seyfu, 1993b, Mulugeta (1978) dealt with some biological aspects of the species. Hailu *et al.* (1990) studied the phenotypic and genotype variability in order to estimate heritability and the expected genetic advance of some quantitative characters in tef plant.

In the study of character associations of tef using 16 cultivars indicated the following results (Hailu, 1988). Grain yield showed positive and significant correlation with kernel weight per main panicle. He noted that using kernel weight per main panicle as a selection criteria would be most useful for grain yield improvement in tef cultivars. Number of seeds per main panicle, productive tillers and 100-kernel weight have also exhibited positive correlation with grain yield though the magnitudes were low.

Murphy (1968) studied different soils in the accessible parts of Ethiopia, and in some instance, their relationships with cereals. As far as the effect of environmental factors on the growth of tef was

concerned, only broad generalizations were made which indicated that the crop is adapted to a great diversity of climatic conditions and soil types (Tadesse, 1969). Sirak (1982), in his study of the effect of different soils on the growth and yield of some tef cultivars, using multiple correlation, found that the tested cultivars adapted better in soils dominated by silt along with adequate quantities of available nitrogen and potassium.

Different source have also given different views on the optimum altitudes for tef production. Some sources state that the best performance occurs at altitude between 1700 - 2400 meters (FAO, 1988). Alkamper(1984), in his study on the influence of altitude on yield of cereals including tef in Ethiopian highlands found that altitude range between 1800 - 2500 meters is suitable for tef production.

Temperature requirements of tef under controlled experiment have been established by Mulugeta (1978) who found that the optimum temperature is between 15^0 - 21^0 c. Some writers have indicated that high temperatures do not affect the growth of tef (Tsegazeab 1974; Brow & Cocheme 1969; FAO 1988). However, according to experiences gained so far from national yield trails, conducted at different locations across the country, tef performs excellently at a temperature range of 10^0 - 27^0 c (Seyfu, 1993).

So far, exact moisture requirement of tef has not been determined on experimental basis. Most of the analyses on the optimum growth of tef dealt with altitude range. According to FAO (1988), the Central Highlands of Ethiopia where most of the tef is grown, get an average annual rainfall of 950 -

1000 mm. The extreme can reach as high as 2500 mm the lower amount can be much lower than the one given above. According to Seyfu (1993) an annual rainfall of 750 - 850 mm and during growing season rainfall of 450 - 550 mm is required for the best growth of tef.

Multivariate studies with special reference to tef

The only study available on tef which uses multivariate techniques is one which was undertaken by Endeshaw (1996). He used principal component analysis in his study of morphological analysis of Eragrostis tef. He found that the effect of natural selection on total variation of characters in one region is different from other regions due to differences in the type and intensity of selection factors involved. However, he has not considered the effect of environmental effects on this cereal.

CHAPTER THREE

METHODS OF STUDY

3.1. The Study Areas and Sources of Data

The data used in this study were collected from Debre Zeit Agricultural Research Center (DZARC) of the Alemaya University of Agriculture, Debre Zeit. Field trials were conducted by the Center at five locations: Debre Zeit black soil; Debre Zeit light soil; Akaki; Koka and Alem-tena in 1991/92, 1992/93 and 1993/94 crop seasons using a complete block design with four replicates. However, Alem-tena was excluded from the analysis due to its incomplete climatic data.

Debre Zeit is located at a latitude of 8°44'N and a longitude of 38°58'E. It has an altitude of 1900 meters above sea level (masl). Debre Zeit has a bimodal rainfall pattern; the main rain season extending from late June to December and the *belg* season extending from early March to May. July and August are the two months which account for over 60% of the rainfall received in the every rainy season.

Akaki is the other substation of the DZARC. It is situated at a latitude of 8°52'N and a longitude of 38°47'E. It has an altitude of 2200 meters above sea level. It is selected mainly for its contrasting environment representing higher altitude areas compared to Debre Zeit. It receives the highest rainfall in August. The main rain season is from mid June to late September which accounts for more than 70% of the total annual rain fall..

Koka is also one of the research stations of the DZARC. It is located at an altitude of 1595 masl. Unlike Debre Zeit, Koka a unimodal rainfall and a short main rainy season extending from June to October.

The data collected were on grain yield and component characters (plant variables) of tef genotypes, climatic variables and soil characters. Climatic data had been also taken from National Meteorological Services Agency to cross check with the record of the Center.

Each location-by-year combination was treated as a separate environment so that a total of 12 environments was used in the analyses. Rainfall and temperature data were all monthly averages, hence a single record of them pertained to four plants in each environment. The total of the thirteen test genotypes in this multi-location trial were sown on 4 m^2 (2m x 2m) field plots at a spacing of 2 m between blocks and 1.5 m between plots within blocks, respectively. However, since one genotype was local check (land-race from each test locality) it was excluded from the analyses. The genotypes of tef codes started from 01 and went up to 13 (Table 3.1). These were the codes given to genotypes of National Yield Trial (NYT).

Table 3.1. NYT codes and Genotype codes of tef used in the study

No.	NYT [a] code	Genotype code
1	01	DZ-01-354
2	03	DZ-01-D2
3	04	DZ-01-58
4	05	DZ-01-795
5	06	DZ-01-D12
6	07	DZ-01-350
7	08	DZ-Cr-358
8	09	DZ-01-1465
9	10	DZ-01-1576
10	11	DZ-01-1809
11	12	DZ-01-D9
12	13	DZ-01-974

[a] Code 02 was for local check.

3.2. Variables Included in the Study

1. **Plant characters**. The parameters assessed involved days to heading, days to maturity, grain yield, plant height and panicle length. Some parameters were excluded from the analyses because they had not been recorded uniformly for all locations.

Days to heading: the number of days from emergence to heading and is taken when the panicle starts to emerge on 75 percent of the plants.

Days to maturity: the number of days from emergence to 75 percent maturity based on visual observation.

Grain yield: the average grain weight in grams of five randomly selected plants per plot.

Plant height: the distance in centimeters from the base of the stem of the main tiller up to the tip of the panicle at maturity averaged over five randomly selected plants per plot.

Panicle length: the length in centimeter from the node where the first panicle branch starts up to the tip of the main panicle at maturity averaged over five randomly selected plants per plot.

2. **Climatic variables**: the variables included in the study were: altitude; average monthly rainfall and average monthly minimum and maximum temperature. For the same reason mentioned before there were variables excluded from the analysis.

3. **Soil characters**: Few soil characters like textures, organic matter, pH, nitrogen and phosphorus were considered in the analysis. Here also some soil characters were excluded from the analysis due to the same reason mentioned above.

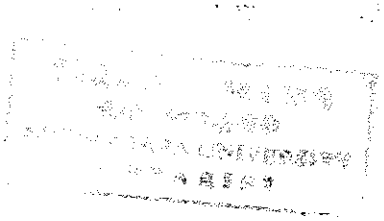
The variables included in the study and their unit of measure are presented in Table 3.2. Table 3.2.

Description of Variables

Variable designation	Definition	Unit of measure
X_1	Days to heading	Days
X_2	Days to maturity	Days
X_3	Grain yield	Gram
X_4	Plant height	Centimeter
X_5	Panicle length	Centimeter
Y_1	Altitude	Meter
Y_2	Average monthly rainfall	Millimeter
Y_3	Average monthly minimum temperature	Degree centigrade
Y_4	Average monthly maximum temperature	Degree centigrade
Y_5	Nitrogen	%
Y_6	Phosphorous	Part per million
Y_7	pH	Unit less
Y_8	Organic matter	%
Y_9	Sand	%
Y_{10}	Silt	%
Y_{11}	Clay	%

3.3 Statistical Procedures Used

To compare the relative magnitude of main effects and interaction variances, variance for yield and component characters due to environment, genotype and genotype by environment effects were estimated using expected mean squares. For this, a combined analysis of variance was used. The ratios of variance components were calculated and presented in a manner similar to that of Peterson et al. (1992). This was followed by canonical correlation analysis (CCA), a multivariate statistical techniques that is used when it is of more interest to observe differences between groups than between individuals. The variability in a large number of variables is firstly reduced to a smaller set of variables that account for most of the variability. The new set of variables, called *canonical variates*, are linear combinations of the original variables. The magnitude of the canonical coefficients indicate their significance in the examination of differences between groups. The canonical correlation analysis was conducted to examine whether a set of environmental variables (soil and climate) could significantly account for the observed variation in yield and component characters of tef genotypes. Lastly, principal components (PCs) analysis was used to determine the underlying source of variability in these traits within and among genotypes. PCs analysis is also a multivariate statistical technique and its main objective is to find a linear transformation of a set of p variates of **X** into a new set denoted by **U**, where the new set has certain desirable properties. These properties, which provide the rationale for using the **U**'s rather than the original **X**'s are: (i) the elements of **U** are uncorrelated with each other (orthogonality); and (ii) each element of **U**, progressing from $\mathbf{u}_1$ to $\mathbf{u}_2$, etc., accounts for as much of the combined variance of the x's as possible, consistent with being



orthogonal to the preceding u's. The mathematical descriptions of the above statistical techniques are given below.

3.3.1. Analysis of Variance

Analysis of variance for all tef genotype attributes were calculated considering each location - by - year combination as a separate environment. Genotypes and environments were treated as fixed effects. Then the combined analysis of variance was done following the procedures described by Gomez and Gomez (1984).

The Model

The response Y_{ijr} of the r^{th} replicate of the i^{th} tef genotype in the j^{th} environment for randomized block design is represented by the model

$Y_{ijr} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijr}$

for i = 1,2, . . ., 12, j = 1,2, . . ., 12 and r = 1, 2, 3, 4,

where

μ = grand mean of overall replicates, genotypes and environments;

α_i = effect of the i^{th} genotype; averaged over the twelve environments, the i^{th} genotype adds α_i to the grand mean μ;

β_j = effect of the j^{th} environment;

$(\alpha\beta)_{ij}$ = genotype - environment interaction of i^{th} genotype in the j^{th} environment;

ε_{ijr} = residual variation contributed by r^{th} replicate of the i^{th} genotype in the j^{th} environment.

The term α_i and β_j are called main effects. They are average effects for each type of genotype and environment. The term $(\alpha\beta)_{ij}$ is an interaction effect.

The following assumptions are important for analysis of variance.

1. Responses to the ijr^{th} treatment combination, Y_{ijr}, are from a normal distribution. This assumption was tested by using the Shapiro - Wilks test (Shapiro & Wilk, 1965).
2. The variances or error mean squares are homogenous. Homogeneity test of variances was conducted using Bartlett's test (Gomez & Gomez, 1984).

The step-by step procedures for the analysis of data combined over environments are:

Step 1. Construction of the outline of the combined analysis of variance over environments, based on randomized block (RCB) design (Table 3.3).

Step 2. Computation of the various sum of squares (SS) in the combined analysis of variance of step 1. These can be computed either from total or mean values. In this study, the computation had been done based on totals. The computations of the various SS follow the standard procedure as:

$$C.F. = \frac{\left(\sum_{i=1}^{e}\sum_{j=1}^{g}\sum_{r=1}^{4} Y_{ijr}\right)^2}{erg};$$

$$SS_E = \sum_{i=1}^{e}\frac{\left(G_i^2\right)}{rg} - C.F.;$$

$$SS_G = \sum_{j=1}^{g}\frac{T_j^2}{re} - C.F. \quad \text{and}$$

$$SS_{GE} = \sum_{i=1}^{g}\sum_{j=1}^{t}\frac{(GE)_{ij}^2}{r} - C.F. - SS_E - SS_G.$$

where

C.F. = Correction factor, SS_E = Environment sum of squares

SS_G = Genotype sum of squares,

SS_{GE} = Genotype by environment interaction sum of squares,

e = number of environments (=12), r = number of replications (=4),

g= number of genotypes (=12),

G_i is the total of the i^{th} environment, T_j is the total of the j^{th} genotype over all e environments, and $(GE)_{ij}$ is the total of the j^{th} genotype in the i^{th} environment.

Computation of pooled error SS as the sum over e error SS, from the individual analyses of variance. That is,

$$SS_{PE} = \sum_{i=1}^{e} (\text{Error SS})_i$$

where $(\text{Error SS})_i$ is error SS from the analysis of variance of the i^{th} environment.

Step 3. Computation of mean squares: For each source of variation, the mean square is computed by dividing each SS by its degrees of freedom (df).

Step 4. Test for the homogeneity of the variances.

Table 3.3. Outline of the combined analysis of variance over e environments, based on RCB design with g genotypes and r replications.

Source of Variation	df	Sum of Squares	Mean Square	Computed F	Expected Mean Square[b]
Environment	e - 1	ESS	MS_E	$(MS_E)/(MS_{res})$	$\sigma^2 + rg\sigma_E^2$
Genotype (G)	g- 1	GSS	MS_G	$(MS_G)/(MS_{res})$	$\sigma^2 + er\sigma_G^2$
G x E	(e-1)(g - 1)	SS_{GE}	MS_{GE}	$(MS_{GE})/(MS_{res})$	$\sigma^2 + r\sigma_{GE}^2$
Pooled error	e(r-1)(g-1)	SS_{PE}	MS_{res}		σ^2

[a] Environment is considered as a fixed variable.

[b] σ_E^2 = Variance due to environment,
σ_G^2 = Variance due to genotype,
σ_{GE}^2 = Genotype by environment interaction variance,
σ^2 = Error variance.

Estimates of variance due to environment, genotype and genotype-by-environment effects were estimated using expected mean squares to compare the relative magnitude of main effects and interactions.

By equating the observed mean squares in the lines of the analysis of variance to their expected values and solving for the variance components, we get

$\hat{\sigma}^2 = MS_{res}$

$\hat{\sigma}^2_{GE} = (MS_{GE} - MS_{res}) / r$

$\hat{\sigma}^2_{E} = (MS_{E} - MS_{res}) / rg$

$\hat{\sigma}^2_{G} = (MS_{G} - MS_{res}) / re$

as estimated values for variance components of residual, genotype - environment, environment and genotype respectively.

3.3.2. Canonical Correlation Analysis

The method of canonical correlation, a multivariate technique attributed to Hoteling (1936), was used to analyze the contributions of environmental factors, climate and soil to the variation of tef genotype. The method enables one to concisely describe the interrelationships of different characteristics between two subsets of data of the whole multivariate system. The method develops linear combinations of each subset of the data, say

$V_i = a_i' X$ and $W_i = b_i' Y$ for $i = 1, 2, \ldots, m$

where

$\mathbf{X}' = (x_1, x_2, \ldots, x_p)$, $\mathbf{Y}' = (y_1, y_2, \ldots, y_q)$

$m = \min(p,q)$ and

$a_i' = (a_{i1}, a_{i2}, \ldots, a_{ip})$

$b_i' = (b_{i1}, b_{i2}, \ldots, b_{iq})$

are standardized canonical coefficients. The new pair, V_i and W_i., called the i^{th} canonical variables (variates), are determined such that

$$\text{Corr}(V_i, V_j) = \text{Corr}(W_i, W_j) = 0$$

for $i \neq j = 1, 2, \ldots, m$, and the parameter ρ_i called the i^{th} canonical correlation is maximized such that $\rho_1 \geq \rho_2 \geq --- \geq \rho_m$.

The determination of $\rho' = (\rho_1, \rho_{2,---}, \rho_m)$, a_i and b_i is based on the eigen structure analysis of the matrix product **H** where

$$\mathbf{H} = \sum_{11}^{-1}\sum_{12}\sum_{22}^{-1}\sum_{21}$$

where $\sum_{11} = \text{Cov}(X)$, $\sum_{22} = \text{Cov}(Y)$ and $\sum_{12} = \sum_{21} = \text{Cov}(X, Y)$

Further,

$$\rho_i = \left|\sqrt{\mu_i}\right|$$

where μ_i is the i^{th} eigenvalue of **H** and so ρ_i is the absolute square root of μ_i. The unknown weight a_i can be obtained from the relation

$$(\mathbf{H} - \mu_i \mathrm{I})\, a_i = 0$$

and

$$b_i = (\sum_{22}^{-1}\sum_{21} \quad a_i) \Big/ \mu_i .$$

The sign and magnitude of the correlation between the original variables and canonical variate are useful in showing

a) which variables contribute most heavily to a canonical variate and the direction of their effects,

b) the existence of the affinities or contrasts among variables in their relationships with a canonical variate.

The correlations thus contribute towards establishing the nature of relationships which may be present between domains. There are two kinds of variables / canonical variate correlations of interest in this study.

i) Intraset Correlation : These are correlations between canonical variates and observed variables of the same domain.

ii) Interset Correlation : These are correlations between canonical variates of one domain and the observed variables of the other. The square of an interset correlation coefficient specifies the proportion of the variance of a variable which is predictable by a canonical variate of the other domain.

Redundancy Measure (RM)

This is used to determine how much of the variance in one set of variables is accounted for by the other set of variables. Redundancy measure can be computed for each canonical correlation.

Let RM_{V_i/W_i} be the amount of variance in the Y variables that is accounted for by X variables for the i^{th} canonical correlation, ρ_i. As given below, computing the RM is a two - step procedure.

First, the average amount of variance in Y variables that is extracted or accounted for by the canonical variate, V_i is computed, and it is equal to the average of the squared loadings of the Y variables on V_i. That is

$$AV(Y/V_i) = \frac{\sum_{j=1}^{q} LY_{ij}^2}{q}$$

where $AV(Y/V_i)$ is the average variance in Y variables that is accounted for by the canonical variate, V_i, and LY_{ij} is the loading of the j^{th} Y variable on the i^{th} canonical variate. Because ρ_i^2 gives the shared variance between V_i and W_i, the redundancy measure is equal to the product of the average variance and the shared variance. That is

$$RM_{V_i/W_i} = AV(Y/V_i) \times \rho_i^2$$

The total variance explained in one set of variables by the other set of variables is called the total redundancy. It is given by

$$RM_{Y/X} = \sum_{i=1}^{m} RM_{Vi/Wi}$$

where $RM_{Y/X}$ is the total redundancy of the Y variables.

Statistical Significance Tests for the canonical correlations

Before interpreting the canonical variates and the canonical correlations, one needs to determine if the canonical correlations are statistically significant. The null and alternative hypotheses for assessing the statistical significance of the canonical correlations are

$H_o: \rho_1 = \rho_2 = \cdots = \rho_m = 0$

$H_a:$ H_o is not true.

The null hypothesis, which states that all the canonical correlations are equal to zero, implies that the correlation matrix containing the correlations among the X and Y variables is equal to zero, i.e., $\mathbf{R}_{xy} = \mathbf{0}$ where $\mathbf{R}_{xy}$ is the correlation matrix containing the correlations between the X and Y variables. This can be tested by using Wilks' λ. Wilks' λ is given by

$$\lambda = \prod_{i=1}^{m}\left(1 - \rho_i^2\right)$$

The statistical significance of the Wilks' λ or the likelihood ratio can be tested by computing the following test statistic :

$$B = -\left[n - 1 - \frac{1}{2}(p + q + 1)\right]\ln\lambda,$$

which has an approximate χ^2 distribution with pq degrees of freedom, for large n (number of observations),

The preceding statistical test is an overall test because it tests for the statistical significance of all the canonical correlations. Rejection of the null hypothesis implies that at least the first canonical correlation is statistically significant. It is quite possible that the remaining m-1 canonical correlations may not be statistically significant. Since the canonical correlations are ordered from the largest to the

smallest. The statistical significance of the second canonical correlation can be tested by computing Wilks' λ after removing the effect of the first pair of canonical variates. In general, the significance of the k^{th} canonical correlation can be assessed by computing the Wilks' λ_k from the modified equation

$$\lambda_k = \prod_{i=k+1}^{m}\left(1-\rho_i^2\right).$$

The value of the corresponding test statistic is equal to

$$B_k = -\left[n-1-\frac{1}{2}(p+q+1)\right]\ln\lambda_k$$

and has an approximate χ^2 with (p-k) (q-k) degrees of freedom.

3.3.3. Principal Component Analysis

Principal components (PCs) analysis is a technique for forming new variables which are linear combinations of the original variables. The maximum number of new variables that can be formed is equal to the number of original variables, and the new variables are uncorrelated among themselves. PCs analysis is conducted either using the covariance matrix $\sum$, or correlation matrix ρ of **X.** Since the units of measurements of the variables in this study were different, the PCs analysis was based on correlation matrix , ρ of **X**. Thus, the PCs are given by

$$\mathbf{U} = \beta' \mathbf{Z} \qquad (1)$$

where $\mathbf{Z}' = (Z_1, Z_2, ..., Z_p)$, Z_i's are the standardized variables of the original variable **X** with $E (\mathbf{Z}) = \mathbf{0}$ and $Cov (\mathbf{Z}) = \rho$ and β is a vector of weights to form the linear combination of the original variables.

Let

$$\rho\,\beta_i = \lambda_i \beta_i \quad (i=1, 2, \ldots, p),$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ are the eigenvalues of ρ, and β_i are the associated eigenvectors normalized by $\beta_i'\beta_i = 1$, ($\beta' = (\beta_1, \beta_2, \ldots, \beta_p)$)and let $U_i = \beta_i' Z$ denote the i^{th} principal component of the standardized variables Z. Thus , the eigenvectors give the weights that are used in the equation (i.e., the principal component) to compute the new variables. That is, the PC analysis reduces to finding the eigenstructure of the correlation matrix of the original data.

The variance of the new variables (U) is given by

$$E(UU') = E\left(\beta' ZZ' \beta\right) = \beta' E\left(ZZ'\right)\beta$$

$$= \beta' \rho \beta \qquad (2)$$

since $E(U) = E\left(\beta' Z\right) = \beta' E(Z) = 0$.

The problem now reduced to finding the weight vector, β', such that the variance, $\beta'\rho\beta$, of the new variable is maximum over the class of linear combinations that can be formed subject to the constraint $\beta'\beta = 1$.

The solution to the maximization problem can be obtained as follows:

Let

$$\phi = \beta' \rho \beta - \lambda(\beta' \beta - 1) \qquad (3)$$

where λ is the Lagrange multiplier. The p - component vector of the partial derivative is given as

$$\frac{\partial \phi}{\partial \beta} = 2\rho\beta - 2\lambda\beta \qquad (4)$$

Setting the above vector of partial derivatives to zero results in the final solution. That is,

$$(\rho - \lambda I)\beta = 0 \qquad (5)$$

For the above system of homogeneous equations to have a trivial solution the determinant of $(\rho - \lambda I)$ should be zero. That is,

$$|\rho - \lambda I| = 0 \qquad (6)$$

Equation (6) is a polynomial in λ of order p, and therefore has p roots. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ be the roots. That is, equation (6) results in p values for λ, and each value is called the eigenvalue or root of the ρ matrix. Each value of λ results in a set of weights given by the p-component vector β by solving the following equations:

$$(\rho - \lambda I)\beta = 0$$

$$\beta' \beta = 1 \qquad (7)$$

Therefore, the first eigenvector, β_1, corresponding to the first eigenvalue, λ_1, is obtained by solving equations

$$(\rho - \lambda_1 I)\beta_1 = 0 \qquad (8)$$

$$\beta'_1 \beta_1 = 1 \qquad (9)$$

Premultiplying equation (8) by β'_1 gives

$$\beta_1'(\rho - \lambda_1 I)\beta_1 = 0$$

$$\beta_1'\rho\beta_1 = \lambda_1\beta_1'\beta_1$$

$$\beta_1'\rho\beta_1 = \lambda_1 \qquad (10)$$

as $\beta_1'\beta_1 = 1$. The left hand side of equation (10) is the variance of the new variable, U_1, and is equal to the eigenvalue, λ_1. The first principal component, therefore, is given by the eigenvector, β_1, corresponding to the largest eigenvalue, λ_1.

Let β_2 be the second p-component vector of weights to form another linear combination. The next linear combination can be found such that the variance of $\beta_2'Z$ is the maximum subject to the constraints $\beta_1'\beta_2 = 0$ and $\beta_2'\beta_2 = 1$. It can be shown that β_2' is the eigenvector of λ_2, the second largest eigenvalue of ρ. Similarly, it can be shown that the remaining principal components, $\beta_3', \beta_4', \cdots \beta_p'$, are the eigenvectors corresponding to the eigenvalues $\lambda_3, \lambda_4, \cdots, \lambda_p$ of the correlation matrix, ρ.

Loadings

The simple correlations between the original and the new variables are called loadings. Loadings give an indication of the extent to which the original variables are influential or important in forming the new variables. The loadings are obtained by using the following equation:

$$l_{ij} = w_{ij}\sqrt{\lambda_i}$$

where l_{ij} is the loading of the j^{th} variable for the i^{th} principal component, w_{ij} is the weight of the j^{th} variable for the i^{th} principal component, λ_i is the eigenvalue (i.e., the variance) of the i^{th} principal component.

All statistical analyses were performed using the SAS statistical package.

CHAPTER FOUR

RESULTS AND DISCUSSION

Based on the statistical techniques discussed in chapter three, the data have been analyzed and their results presented and discussed as follows.

4.1. Analysis of Variance

Test of Normality

The Shapiro - Wilks test is utilized to test the normality of plant variables. Table A.1 (Appendix) contains this result.

The results in Table A.1 suggest that the Shapiro-Wilks statistic for each variable in each genotypes is not significant. Hence, the distributions of these variables are normal.

Test of Homogeneity of Variances

The results of Chi-square test (Bartlett's test for homogeneity of variance) of twelve error variances for each plant variables X_1, X_2, X_3, X_4, and X_5 are 5.036, 1.503, 6.088, 4.810, and 10.402, respectively .

Since the computed Chi-square values are smaller than the corresponding table value at 5 % level of significant (i.e., $\chi^2(11) = 19.67$), heterogeneity of variance is not indicated. Thus, the hypothesis

that the error variances are homogeneous should not be rejected for all plant variables considered in this study.

Combined Analysis of Variance

Combined analysis of variance was carried out after testing for homogeneity of variance using Bartlett's test. The anova (Table 4.1) showed significant differences between years and locations among the main effects while there was no difference between genotypes. However, There was a significant interaction between locations and genotypes.

Table 4.1. Analysis of variance table for grain yield of tef

Source	DF	Sum of Squares	Mean Square	F values
Replication	3	151481.2	50493.7	0.90
Year (Y)	2	8331849.4	4165924.7	74.3**
Location (L)	3	42089556.9	14029852.3	250.2**
YxL	6	25722768.5	4287128.1	76.5**
Genotype (G)	11	384081.5	34916.5	0.6
YxG	22	1318421.4	59928.2	1.1
LxG	33	2783765.3	84356.5	1.5*
YxLxG	66	3283369.8	49748.0	0.9
Error	429	24052435.3	56066.3	
Total	575	108117729.5		

* significant at 5%, ** significant at 1%

The results of combined analysis of variance of plant variables over 12 environments are summarized in Table A.2 (Appendix). This analysis indicated that there were highly significant (at 1% level of significance) differences among the environments for grain yield and component characters of tef genotypes. This suggests that the locations (Debre Zeit black soil, Debre Zeit light soil, Koka and Akaki) are different from each other and represent different environmental conditions for different crop seasons. Thus, the multi-location testing of diverse tef genotypes for those characters were quite logical for tef improvement program. Highly significant differences were also observed among the genotypes when tested across many locations for component characters showing that they performed differently across a range of environments and non significant difference was observed for yield (Table A.2, Appendix).

The interaction effects of genotype and environment were highly significant for plant height and panicle length. The relative magnitudes for the effects of genotype, environment and their interaction were examined by comparison of variance components (Table 4.2). A large influence of environment relative to that of genotype was noted for all plant variables considered in this study. The variance ratio in Table 4.2 indicates this fact.

Table 4.2. Variance components and their ratios estimated for environment and genotype main effects and their interaction for grain yield and component characters.

Plant variable[a]	Variance Component			Variance ratios[b]	
	Env	Gen	Env x Gen	(Env)/(Gen)	(Gen)/(Env x Gen)
X1	47.18	3.33	0.31	14.17	10.74
X2	300.49	4.95	2.85	57.90	1.82
X3	143162.91	322.12	2664.85	444.43	0.12
X4	171.21	9.68	116.19	17.69	0.08
X5	41.98	3.44	41.32	12.20	0.08

[a] As presented in Table 3.2.
b (Env)/ (Gen) = (Environment)/(Genotype).
(Gen)/ (Env X Gen) = (Genotype)/(Environment x Genotype).

4.2. Canonical Correlation Analysis (CCA)

Relationships between environment variables

The correlation between environment attributes (Y_1, Y_2, Y_3, ..., Y_{11}) were examined through the use of simple linear correlations (Table A.3, Appendix). This indicated that no single parameter of environment group could explain variation in grain yield and component characters of tef genotypes. Most of the attributes in environment group were interrelated which excluded the use of conventional multiple linear regression analysis. Thus, canonical correlation analysis was considered as an appropriate method to explain the observed variation in grain yield and component characters of tef genotypes.

CCA was conducted to examine whether a set of environment variables (soil and climate) could account for the observed variation in tef genotypes. It sets out to find linear combinations of the two

sets of variables, tef genotypes and environment , such that the linear combinations (canonical variates) have maximal correlations.

The canonical correlation coefficients $\rho' = (\rho_1, \rho_2, \rho_3, \rho_4, \rho_5)$ between grain yield and component characters of 12 tef genotypes and set of environment variables were summarized in Table 4.3.

Table 4.3. Canonical correlation coefficients between grain yield and component characters of 12 genotypes of tef and a set of environment variables

Tef Genotype	Canonical Variate 1	2	3	4	5
01	0.873	0.593	0.409	0.164	0.001
03	0.897	0.735	0.359	0.201	0.002
04	0.903	0.645	0.297	0.223	0.003
05	0.939	0.689	0.258	0.174	0.005
06	0.922	0.699	0.349	0.170	0.002
07	0.909	0.675	0.581	0.137	0.026
08	0.932	0.753	0.470	0.148	0.007
09	0.911	0.651	0.308	0.236	0.005
10	0.825	0.514	0.252	0.073	0.037
11	0.894	0.712	0.357	0.236	0.025
12	0.910	0.733	0.374	0.134	0.005
13	0.891	0.567	0.185	0.095	0.021

Before interpreting the canonical variates and canonical correlations it is necessary to determine if the canonical correlations are statistically significant. Therefore, this was done first. Table A.4 (Appendix) contains a test of individual canonical correlations. The results showed that the first canonical correlation of all tef genotypes, and the second canonical correlation of ten genotypes (excluding genotypes 10 and 13) were statistically significant at 5 % level of significance.

Table A.5 (Appendix) presents the percentage contribution of the individual canonical variate for the five components while Table A.6 (Appendix) gives the correlation (loading) between the individual variables of the two sets and the significant canonical variates of environment groups together with, redundancy measures.

Based on χ^2 - test statistic, a test of the null hypothesis:

$$H_o: \sum_{12} = 0$$

i.e., the absence of overall correlations between the two subsets of data is given in Table A.4 (Appendix).

The results of this test led to the rejection of the overall hypothesis of no association, hence the two subsets of data on tef genotype and environment variables can be considered to be linearly related in twelve tef genotypes. Furthermore, among the tests on the joint nullity of the canonical correlations groups, only the groups $(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5)$ and $(\rho_2, \rho_3, \rho_4, \rho_5)$ were significant for all genotypes, except genotypes 10 and 13 , whereas only $(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5)$ showed strong evidence to reject the null hypothesis in the case of genotypes 10 and 13. Thus, only the first canonical correlation

coefficient of genotypes 10 and 13 and the first two canonical correlations of the remaining genotypes were statistically significant. It may therefore be inferred that the environmental variables and the genotypes attributes of tef are linearly related and that the major part of the relationship can be accounted for by the first pair of canonical variate (V_1, W_1) for genotypes 10 and 13, and by the first two pairs of canonical variates (V_1, W_1) and (V_2, W_2) for other genotypes.

Hence, the correlations between the two sets of variables can be accounted for by just the first pair of canonical variates in the case of genotypes 10 and 13. This canonical pair accounted for 75.73 and 83.13 percent of the total correlation between the set of genotype variables and the set of environment variables, respectively. Thus, if we concentrate only on the first pair of canonical variate for genotypes 10 and 13 then we shall forsake only 24.27 and 16.87 percent, respectively of the information on the total correlation.

Similarly, the first two canonical pairs of genotype and environment variables of genotypes 01, 03, 04, 05, 06, 07, 08, 09, 11 and 12 cumulatively accounted for 90.36, 93.40, 93.96, 96.93, 95.03, 88.50, 94.00, 94.09, 92.49 and 94.43 percent, respectively (Table A.5, Appendix).

Having assessed the statistical significance of the canonical correlations, the next step is to interpret the canonical variates. Generally, only those canonical variates whose canonical correlations are statistically significant are interpreted.

A ± 0.5 and above intraset and interset correlation coefficients were used for interpretation.

The results in Table A.6 (Appendix) shows that, on the environment domain, except altitude and clay content, all environment variables contribute to the canonical variate CV_1 in the same direction for most of the genotypes. CV_1 was characterized particularly by altitude, minimum and maximum temperature, nitrogen, silt and clay for genotypes 01, 03, 04, 05, 06, 07, 08, 09, 12 and 13, for genotype 10 by altitude, minimum and maximum temperature, silt and clay and for genotype 11 by altitude, minimum and maximum temperature, nitrogen, sand and clay. In genotypes 03, 09, and 11 the contribution of sand to variance within this domain was relatively high.

The large positive loading of altitude and CV_1 may indicate that the variation in the traits of tef genotypes highly depends on altitude since the contribution of this variable to the variation accounted by CV_1 is high for all tef genotypes. The large negative intraset correlations of temperature and CV_1 my indicate the negative effect on the traits of tef genotypes when the level of temperature increases. High temperatures facilitate rapid evaporation of water from soil. As a result, soils become dry. Consequently, the growth of both roots and shoots of the genotypes is reduced. So the number of days to heading and days to maturity decreases, plant height and panicle length shortened and grain yield reduced (note that the interset correlation of attributes with CV_1.). However, the joint effect of moderate altitude (1800 - 2100 m), temperature and adequate rainfall during the growing season on tef performances is positive (Seyfu, 1993).

High negative intraset correlation of nitrogen with CV_1 probably indicated that the excessive of nitrogen might delay crop maturity by encouraging excessive biomass growth of the tef plants.

Similarly, the large negative intraset correlation of silt content with CV_1 confirms the fact that soils dominated by silt have small pore spaces which slow down water movement and characterize by high formation; hence root growth may be restricted and the soil may become water logged after rain. Therefore, this had its own contribution to the problems observed above in the attributes of tef genotypes.

The effect of high clay content in the soil on the variation among genotypes growth is indicated by the positive intraset correlation of clay with CV_1. Soils with high clay content generally have high water-holding capacity but inadequate aeration. The decreased aeration resulting from soil compaction causes a decrease in the uptake of potassium (Thompson et. al., 1978). Potassium is a macronutrient needed in relatively large amounts by all plants and is important in metabolism in the formation of carbohydrates and proteins; it also aids in the uptake of other nutrients and in the movement within the plant (Thompson et. al., 1978). The water holding capacity of the soils might have contributed to the positive interaset correlation.

The second variate CV_2 is the difference between the total contribution of altitude, rainfall, nitrogen, phosphorous, organic matter and sand as one group, and that of minimum and maximum temperature, silt and clay as the other group. Large positive values of intraset correlations were observed for rainfall, nitrogen, phosphorous and organic matter with CV_2. It, therefore seems that the nitrogen, phosphorous and organic matter of the soils have been the edaphic factors that have high contribution to grain yield (note the interset correlation of grain yield with CV_2).

From the interset correlations (Loadings) between grain yield and CV_2 of genotypes 03, 04, 05, 06, 07, 08, 10, 11, and 12, it seems likely that these genotypes have better performance in terms of grain yield on soils dominated by nitrogen, phosphorous, organic matter and to some extent, by sand. As in CV_1, the effect of pH is negligible in CV_2 .

The variate CV_1 extracts about 28, 32, 29, 32, 31, 31, 32, 32, 31, 29, 29 and 32 percent of the variance in the environment groups for genotypes 01, 03, 04, 05, 06, 07, 09, 10, 11, 12 and 13, respectively. The variate CV_2 explained approximately an additional amount of 25, 22, 20, 22, 25, 22, 19, 16, 21, 27, 26 and 25 percent of the variation in the environment domain for genotypes 01, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12 and 13, respectively. Thus, jointly CV_1 and CV_2 extracted about 53, 54, 49, 54, 56, 53, 54, 51, 47, 50, 56 and 57 percent of the total variance in the environment domain (Table A.6, Appendix), the remaining being attributed to CV_3, CV_4 and CV_5 .

The redundancy measure of the five attributes of genotypes, 0.401, 0.439, 0.488, 0.498, 0.507, 0.447, 0.522, 0.466, 0.282, 0.401, 0.475 and 0.417 for the first canonical variate, CV_1 indicate that about 40, 44, 49, 49, 51, 45, 52, 47, 28, 40, 48 and 42 percent of the variation in grain yield and component characters within genotypes 01, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12 and 13, respectively are accounted for by the environment variables.

The total redundancy for yield and plant characters of genotypes 01, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12 and 13 are 0.590, 0.633, 0.632, 0.679, 0.691, 0.601, 0.728, 0.668, 0.492, 0.596, 0.668 and 0.564, respectively. That is, the environment variables considered in this study jointly account for about 59, 63, 63, 68, 69, 60, 71, 67, 49, 60, 67, and 56 percent of the within variation in the five attributes of genotypes 01, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12 and 13, respectively. Hence the

remaining variation in these attributes might be due to other plant variables of the genotypes and other environment variables not included in this study.

4.3. Principal Component Analysis

The relationship among the grain yield and component characters as determined over 12 environments for 12 tef genotypes were evaluated by principal component analysis (using correlation matrix) to determine the source of underlying variability within genotype and among genotypes. A loading coefficient, which is a measure of association between the attributes and PC, of ± 0.5 and higher values were used for interpretation.

Table A.7 (Appendix) presents the eigenvalue, percentage contribution of the individual principal components and loading coefficients between PCs and the plant variables. The principal components (PCs) with eigenvalue greater than one are retained for interpretation. That is, using the rule of eigenvalue greater - one.

Genotype 01

The first two principal components (PCs) accounted for 81.8 % of the variability of yield and component characters of genotype 01 (Table A.7, Appendix). The magnitude of the loading coefficients indicated that all the plant variables have high contributions to PCs, and this is true for all the genotypes. The first PC accounted for 50.4 % of the variability and was related to days to heading, days to maturity, plant height and panicle length. As these attributes are component

characters, the PC_1 therefore reflects the component characters variation within this genotype. That is, PC_1 is a measure of the component characters variation across this genotype. The second PC accounted for 31.4 % of the variability and high loading coefficients were obtained for days to heading, days to maturity and grain yield.

Genotype 03

The first two PCs jointly accounted for 80.1 % of the total variation in yield and component characters (Table A.7, Appendix). For a given the loading coefficient, PC_1 appears to be a contrast of a weighted average of days to heading and days to maturity versus a weighted average of plant height and panicle length, whereas the PC_2 appears to be a weighted average of days to maturity and yield.

Like genotype 01, it might be possible to say that PC_1 reflects the component characters variation within genotype 03. Grain yield has the highest contribution to the variation explained by PC_2.

Genotype 04

The first two PCs collectively accounted for 81 % of the total variation in yield and component characters of this genotype(Table 7, Appendix). Except that PC_1 was the contrast of a weighted average of yield, plant height and panicle length versus a weighted average of days to heading and days to maturity, PC_2 of genotype 04 had similar nature as that of genotype 03.

Genotype 05

PC_1 and PC_2 jointly accounted for 84.1 of the variability of yield and component characters of genotype 05 (Table A.7, Appendix). PC_1 and PC_2 accounted for 59.7 and 24.4 % of the total variation within genotype 05, respectively. The nature of the loading coefficients in PC_1 and PC_2 are similar to the previous genotypes.

Here also it might be possible to say that the first PC reflects a component character variation within genotype 05. The comparison of loading coefficients in PC_2 indicated that yield had the highest contribution to the variation explained by PC_2.

Genotype 06

The first two PCs collectively accounted for 85.6 % of the total variation in yield and component characters of this genotype (Table A.7, Appendix). The first PC (PC_1) represented 56.9 % of the variation within genotype 06, and dominated by days to heading, days to maturity, plant height and panicle length. The second PC (PC_2), which was independent of PC_1, accounted for 28.7 % of the variation within genotype 06. PC_2 had the strongest relation with yield, followed by days to maturity, days to heading, plant height and panicle length.

Genotype 07

The first PC, which explained 52. 6 % of the total variation of the five attributes of genotype 07, included all component character. Like genotypes 01, 03, 05, 06, it might be called the component characters variation. The second PC accounted for 30.4 % of the variation and was related to days to heading, days to maturity and yield.

Genotype 08

The first PC accounted for 60.4 % of the variance in yield and component characters and all the component characters had high loading on the first PC. Therefore, like genotypes 01,03,05,06 and 07, the first PC probably represents the component characters variability within genotype 08.
The second PC had a high loading for yield, so PC_2 likely represents the variation explained by this variable alone. The percentage contribution of PC_2 was 31.1 %.

Genotype 09

The first two PCs, PC_1 and PC_2, collectively accounted for 83.3 % of the total variation in yield and component characters (Table A.7, Appendix). Of the total variation 52.2 % was explained by PC_1 which showed similar property with respect to its loading coefficients, like the above genotypes (except genotype 04).

The percentage contribution of PC_2 was 31.1 % and the highest loading of PC_2 was for days to maturity, followed by days to heading and yield. The loadings for plant height and panicle length with PC_2 were low like the previous genotypes.

Genotype 10

Unlike the previous genotypes, the first PC of this genotype had relation only with yield, plant height and panicle length. The percentage contribution explained by PC_1 was 41.5 %. The second PC accounted for 31.1 % of the variability and like genotypes 01 and 07, high loading coefficients were obtained for days to heading, days to maturity and yield. However, the contributions of plant height and panicle length to PC_2 were almost negligible.

Genotype 11

The first two PCs, PC_1 and PC_2, jointly accounted for 82 % of the variability in yield and component characters. Unlike that of the previous genotypes, PC_1 was dominated by days to heading, plant height and panicle length and accounted for 47.3 % of the total variability. The second PC accounted for 34.7 %, which was the largest of all of the percentage corresponding to PC_2, observed in 12 tef genotypes and high loading coefficients were obtained for days to heading, days to maturity and yield.

Genotype 12

The first PC of this genotype was a contrast of a weighted average of yield, plant height and panicle length versus a weighted average of days to heading and days to maturity just as that of genotype 04. PC_1 accounted for 54.8 % of the variability in yield and component characters. However, the second PC was a weighted average of days to maturity and yield and accounted for 26.4 % of the total variance of yield and component characters.

Genotype 13

As in the case of genotypes 04 and 12, the first PC appears to be the contrast of a weighted average of yield, plant height and panicle length versus a weighted average of days to heading and days to maturity. It accounted for 49 % of the variability in yield and component characters of genotype 13. The second PC accounted for 33 % of the variability of these attributes and the highest loading was observed for days to heading, followed by days to maturity, yield, plant height and panicle length.

PC analysis among Genotypes

PCs analysis was also used to determine the underlying source of variation among 12 tef genotypes. The first two PCs accounted for 82.2 of the variation in yield and component characters observed in 12 tef genotypes (Table 4.3). The magnitude of loading coefficients indicated that like the within

genotype case, all attributes of genotypes used in this study had their own contribution to the total variability among genotypes. The first PC accounted for 62.5 % of the variability and was highly related to days to heading, days to maturity, plant height and panicle length. As these attributes are all component characters, the PC_1 thus reflects the component character differences among tef genotypes. Thus the variation among tef genotypes more accounted for by the component characters. The second PC accounted for 19.7 % of the variability and high loading coefficient was obtained for yield therefore PC_2 measures the yield differences among genotypes.

Table 4.4 Principal component analysis of yield and component characters data among genotypes

Parameter		Principal Component				
		PC_1	PC_2	PC_3	PC_4	PC_5
Eigenvalue		3.12	1.00	0.59	0.21	0.08
% variation explained		62.5	19.7	11.8	4.3	1.7
Loading	X_1	0.88	-0.27	0.15	0.35	0.08
	X_2	0.78	0.01	-0.61	-0.09	0.10
	X_3	0.28	0.98	0.08	0.08	0.04
	X_4	0.85	-0.07	0.43	-0.28	0.08
	X_5	0.97	0.03	-0.05	-0.01	-0.24

Because of the non availability of the proper data the environments have been taken as location x year. An attempt has been made regarding yield data to split environments considered into separate factors, that is year, location and their interactions, considering also their effect on genotypes. It has been observed that significance of separate factors cannot be ignored. It is recommended that for such research can be undertaken in the light of results obtained for the yield data.

Multivariate analysis techniques are also highly recommended in the study of relationships existing between environmental and genotypic variables.

Despite all the limitations of this study, I am confident that the results would form a useful basis for tef improvement program in the country.

by the environment domain considered in this study. The remaining might be accounted for by the inherent nature of tef genotypes or environmental factors not included here.

Principal components analysis indicated that each of the genotype variables had their own contribution to the variation observed in genotypes and among all tef genotypes. Moreover, the component characters of tef genotypes explained most of the variation within genotypes (for most genotypes) and among genotypes.

With the present analysis, it was believed that a general picture of environmental effects on tef genotype had been obtained. However, the writer strongly recommends that this study should be followed by further research by tef breeders, to cover more genotypes and environments based on the above conclusions. Researches should also be designed to determine the optimal levels of the environmental variables, for instance soil moisture stress and excess at different phenological stages of tef development. Such analysis will facilitate acquisition of knowledge on the water requirement of tef and after which to recommend irrigation whenever there is shortage, and drainage when there is excess moisture.

There is ample information pertinent to this work at Debre Zeit Agricultural Research Centre. However, some of the information which has not been not been consistently and uniformly recorded has been excluded from the analysis. It is recommended that future studies in tef research record as many parameters as consistently as possible so that a deeper understanding of the interrelationships existing between environmental and genotypic variables are acquired through statistical analyses.

REFERENCES

Alkamper, J. 1984. The influence of altitude on crop growing and yield. *Plant Research and Development*. **20**, pp. 59 - 71.

Byth, D.E., Eismann, R.L. & De Lacy, I.H. 1976. Two-way pattern analysis of a large data set to evaluate genotypic adaptation. *Heredity* **37**, pp. 215-230.

Brown, L.H. and Cocheme J. 1969. A study of the agroclimatology report, FAO/ UNESCO / WMO Interagency, Agroclimatology project, FAO, Rome.

Central Statistical Authority. 1996. Agricultural sample survey 1995/96 (1988, E.C.), Report on Area and production for major crops. *Statistical Bulletin* **152**, Addis Ababa, Ethiopia.

Endashaw Bekele. 1996. Morphological analysis of Eragrostis Tef: Detection for regional patterns of variations. *SINET: Ethiop. J. Sci.* **19(1)**: pp. 117-140.

FAO. 1988. Traditional food plants, FAO, Food and Nuitrition paper. **No. 42**, FAO, Rome.

Freeman, G.H. and Dowker, B. D., 1973. The analysis of variation between and within genotypes and environments. *Heredity*, **30**, pp. 97 - 109.

Gomez, A.K. & Gomez, A. 1984. *Statistical procedures for agricultural research.* 2nd Ed. A wiley-interscience Publication, New York.

Graybosch, R., Peterson, C.J., Moore, K.J., Stearns, M & Grant, D.L. 1993. Comparative effects of wheat flour protein, liquid, and pentosan composition in relation to baking and milling quality. *Cereal Chem.* **70**, pp. 95 - 101.

Hailu Teffera, 1988. Variability and association of characters in Tef [*Eragrostis tef* (zucc.) Trotter] Cultivars. M.Sc. Thesis, AUA, Dire Dawa, Ethiopia.

Hailu Tefwra, Seyfu Ketema and Tesfaye Tesemma, 1990. Variability, heritability and genetic advance in tef (*Eragrostis tef* (zucc.) Trotter) cultivars. *Trop. Agric.* **67** (**4**), pp. 317 - 320.

Hailu Teffera, Mulu Ayele & Kebebw Assefa, 1995. Improved varieties of Tef (*Eragrostis tef*) in Ethiopia releases of 1970 - 1995. *Res. Bulletiny* **1**, DZARC, AUA, Dire Dawa, Ethiopia.

Hofstra, G., 1972. Responses of soybean to temperature under high light intensities. *Can. J. Plant Sci.* **52**, pp 535-543.

Hotelling, H. 1936. Relations between two sets of variates. *Biometrica*, **28**, pp 321-377

Jennings, D.L., 1979. The occurrence of multiple fruiting laterals at single nodes of raspberry Canes. *New Phytol.* **82**, 365 - 374.

Jenson, R.J. & J.F. Hancock, 1982. Multivariate relationships among California strawberries. *Bul. Torrey Bot. Club* **109**, 136 - 147.

Johnson, R.A. & Wichern, D.W., 1992. *Applied multivariate statistical analysis.* 3rd Ed., New Jersey: Prentice Hall.

Katiyar, R.P., 1979. Correlations and path analysis of components in chickpea. *Indian J. Agric. Sci.* **49 (1)**, 35 - 38.

Kempton, R. E. 1984. The use of biplots in interpreting variety by envrionment interactions. *J. Agric.Sci.* (Cambs) **103**, pp 123-135.

Lezzoni, A.F. & Pritts, M.P., 1991. Applications of principal component analysis to horticultural research. *Hort. Sci.* **26**, 334 - 338.

Melak-Hail Mengesha, 1964. *Eragrostis tef* (zucc.) Trotter; its embryosac development, genetic variability and breeding behaviour. Ph. D. Thesis, Purdue Univ., Lafayette. U.S.A. pp. 83.

Melak-Hail Mengesha, 1966. Chemical composition of tef (Eragrostis tef) compared with that of wheat, barley and grain sorghum. *Econ. Bot.*, **20**, 268 - 273.

Murphy, M.F., 1968. A report on the fertility status and other data on some soil of Ethiopia. *Exper. Sta. Bull.* **No. 44**, HISU. College of Agri. Dire Dawa, Ethiopia.

Nevo, E., D. Zahary, A. Brown, & H. Habe, 1979. Genetic diversity and environmental associations of wild barley. Hordeum spontaneum, in Israel. *Evolution* **33**, 815 - 833.

Oranga, H.M., 1990. Canonical analysis of the role of nutrition on the resistance of indigenous cattle to ticks. *Proceedings of the first scientific meeting of the Biometrics Society. Biom. for dev.* Kenya group and East.

Osborne, B.G., Salmon, S & Curtis, P.S., 1993. Quality assessment of new wheat varieties. In J.R.N. Taylor, P.G. Randall & J.H. Viljoen (eds.). *Selected papers from the ICC International Symposium*, pp. 319 - 333. CSIR Publishers, Pretoria.

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

Pimentel, Richard A., 1979. Morphometrics, the multivariate analysis of biological data. California Polytechnic State University. San Luis Obispo, California.

Preston, K.R., Lukow, O.M. & Morgan, B., 1992. Analysis of relationships between flour quality properties and protein fractions in a world wheat collection. *Cereal Chem.* **69**, 560 - 567.

Rouk, H.F., & Melak-Hail Mengesha, 1963. An introduction to tef (*Eragrostis abyssinica Schrad.*) a nutritious cereal grain of Ethiopia. *DZARC Bull.* **No. 26**, AUA, Dire Dawa, Ethiopia.

Seyfu Ketema, 1988. Tef [Eragrostis tef (zucc.) Trotter]: Basics of its taxonomy, morphology, cytology and breeding. In first annual conference of the crop science committee of Ethiopia. Addis Ababa, Feb. 8 - 9.

Seyfu Ketema, 1993b. Tef (Eragrostis tef) Breeding, genetic resources, utilization and role in Ethiopia agriculture. IAR, Addis Ababa, Ethiopia.

Shapiro, S.S. & Wilk, M.B., 1965. An analysis of variance test for normality (Complete Samples). *Biometrika* **52**, 3 and 4, pp. 591 - 611.

Sirak Isak, 1982. The effect of different soils on the growth and yield of some tef cultivars. M.Sc. Thesis, AAU. Addis Ababa, Ethiopia.

Subhash Sharma, 1996. *Applied multivariate techniques*. John Wiley & Sons, Inc. New York.

Tadesse Ebba, 1969. Tef (Eragrostis tef). The cultivation, usage, and some of the known diseases and insect pests. Part I. *Agri. Expt. Sta. Bull.* No. 60. Haile Selassie I University, College of Agriculture, Dire Dawa, Ethiopia.

Tadesse Ebba, 1975. Tef (Eragrostis tef) cultivar morphology and classification. Part II. *Agri. Expt. Sta. Bull.* **No. 69**. AAU, College of Agriculture, Dire Dawa, Ethiopia.

Tareke Berhe, 1972. Future properties of mutation breeding on the improvement of Eragrostis tef (zucc.) Trotter: Paper presented in the joint FAO / IAEA meeting for protein improvement held in Munich, Germany, 26 - 30, June 1972.

Tsegazeab Wolde, 1974. Agroclimatology of teff. In Agroclimatology of the highgl;ands of eastern Africa. World Meteorological Organaization, No. 389. FAO/UNESCO/WMO Interagency Progect on Agroclimatology, WMO, Geneva

Van Lill, D., Purchase, J.L., Smith, Marie F., Agenbag, G.A. & DE Villiers, O.T., 1995. Multivariate assessment of environmental effects on hard red winter wheat. I. Principal-components analysis of yield and bread-making characteristics. *S. Afr. J. Pl. Soil*, **vol. 12 (4)**, 158 - 163.

Van Lill, D., Purchase, J.L., Smith, Marie F., Agenbag, G.A. & DE Villiers, O.T., 1995. Multivariate assessment of environmental effects on hard red winter wheat. II. Canonical correlation and canonical variate analysis of yield, biochemical and bread-making characteristics. *S. Afr. J. Pl. Soil*, **12** (4),164 - 169.

Vavilov Ni, 1951. The origin variation, immunity and breeding of cultivated plants. Ronald press, New York, USA, pp. 37 - 38.

Williams, W.T., C.A.P. Boundy, & A.J. Millington, 1977. The effect of sowing date on the growth and yield of three sorghum cultivars in the Ord River Valley. II. The components of growth and yield. *Austral. J. Agr. Res.* **28**,381 - 387.

APPENDIX

Table A.1. Normality test of grain yield and component characters of 12 tef genotypes

Genotype	Plant Variables[a]	Shapiro - Wilks Statistic	D.F.[b]	Significance
01	X_1	0.9301	48	0.1010
	X_2	0.9604	48	0.2285
	X_3	0.9526	48	0.0889
	X_4	0.9673	48	0.3658
	X_5	0.9659	48	0.3370
03	X_1	0.9271	48	0.1000
	X_2	0.9500	48	0.0710
	X_3	0.9639	48	0.2971
	X_4	0.9807	48	0.0743
	X_5	0.9520	48	0.0863
04	X_1	0.9326	48	0.1034
	X_2	0.9481	48	0.0581
	X_3	0.9355	48	0.0811
	X_4	0.9846	48	0.0860
	X_5	0.9468	48	0.0594

[a] As presented in Table 3.2.

b Degrees of Freedom

Table A.1 (Continued)

Genotype	Plant Variables[a]	Shapiro - Wilks Statistic	D.F.[b]	Significance
05	X_1	0.9304	48	0.1018
	X_2	0.9639	48	0.2984
	X_3	0.9549	48	0.1177
	X_4	0.9316	48	0.1033
	X_5	0.9344	48	0.1067
06	X_1	0.9766	48	0.1001
	X_2	0.9535	48	0.0968
	X_3	0.9672	48	0.3635
	X_4	0.9790	48	0.1000
	X_5	0.9426	48	0.1036
07	X_1	0.9306	48	0.1014
	X_2	0.9463	48	0.0679
	X_3	0.9615	48	0.2503
	X_4	0.9377	48	0.1007
	X_5	0.9442	48	0.0519
08	X_1	0.9616	48	0.0936
	X_2	0.9612	48	0.2449
	X_3	0.9658	48	0.3369
	X_4	0.9531	48	0.0938
	X_5	0.9340	48	0.0631

Table A.1 (Continued)

Genotype	Plant Variables[a]	Shapiro - Wilks Statistic	D.F.[b]	Significance
09	X_1	0.9393	48	0.1084
	X_2	0.9480	48	0.0570
	X_3	0.9663	48	0.3468
	X_4	0.9368	48	0.1062
	X_5	0.9667	48	0.0841
11	X_1	0.9623	48	0.2641
	X_2	0.9633	48	0.2865
	X_3	0.9493	48	0.0666
	X_4	0.9402	48	0.2097
	X_5	0.9477	48	0.0548
12	X_1	0.9611	48	0.2132
	X_2	0.9457	48	0.0642
	X_3	0.9636	48	0.2924
	X_4	0.9362	48	0.1027
	X_5	0.9555	48	0.1194
13	X_1	0.9341	48	0.1603
	X_2	0.9353	48	0.1709
	X_3	0.9665	48	0.3497
	X_4	0.9412	48	0.2031
	X_5	0.9647	48	0.3138

Table A.2. Combined analysis of variance for grain yield and component characters in 12 tef genotypes across 12 environments and 4 replications

Plant Variable	Mean Square Values for				Computed F[a]		
	Environment	Genotype	Environment x Genotype	Error	E	G	G x E
Days to heading	2271.54	166.59	8.13	6.88	142.37	24.21	1.18
Days to maturity	14455.01	280.85	43.08	31.67	104.27	8.87	1.36
Grain yield	2944895.19	451594.21	127321.711	36840.77	115.68	12.26	3.46
Plant height	8279.71	526.38	855.69	61.62	33.91	13.89	8.54
Panicle length	2029.63	179.65	218.24	14.36	71.30	15.20	12.51
Degrees of freedom	11	11	121	396			

[a] E= Environment, G= Genotype, G x E = Environment x Genotype

Except for GxE interaction of days to heading and days to maturity all F-ratios are highly significant

Table A.3. Correlation coefficient among environmental variables measured in 1991/92, 1992/93 and 1993/94 crop seasons across four locations.

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10
Y2	0.219									
Y3	-0.985**	-0.325								
Y4	-0.987**	-0.270	0.984**							
Y5	-0.261	0.583**	0.107	0.151						
Y6	0.475**	0.651**	-.597**	-.560**	0.609**					
Y7	-0.173	-0.110	0.188	0.183	0.080	-0.492**				
Y8	0.276	0.666**	-0.414*	-0.373*	0.746**	0.976**	0.465**			
Y9	-0.315	0.249	0.235	0.258	0.651**	-0.032	0.808**	0.074		
Y10	-0.969**	-0.314	0.979**	0.974**	0.134	-0.649**	0.373*	-0.468**	0.391*	
Y11	0.485**	-0.183	-0.410*	-0.430*	-.645**	0.129	-0.783**	-0.009	-0.983**	-0.551**

* $p < 0.01$* $p < 0.001$

Table A.4. Chi-square tests of the joint nullity of the smallest 5 – k canonical correlations

Tef Genotype	Roots	Wilks' λ_k	Test statistic[b] value (B_r)	D.F[c]
01	1,2,3,4,5	0.025	142.022*	55
	2,3,4,5	0.198	62.350*	40
	3,4,5	0.493	27.229	27
	4,5,	0.835	6.942	16
	5	0.999	0.038	7
03	1,2,3,4,5	0.014	164.345*	55
	2,3,4,5	0.135	77.095*	40
	3,4,5	0.511	25.848	27
	4,5,	0.798	8.687	16
	5	0.998	0.077	7
04	1,2,3,4,5	0.029	152.288*	55
	2,3,4,5	0.193	63.335*	40
	3,4,5	0.545	23.368	27
	4,5,	0.774	9.863	16
	5	0.997	0.116	7
05	1,2,3,4,5	0.019	173.629*	55
	2,3,4,5	0.189	64.141*	40
	3,4,5	0.610	19.030	27
	4,5,	0.822	7.546	16
	5	0.995	0.193	7
06	1,2,3,4,5	0.012	170.279*	55
	2,3,4,5	0.162	70.076*	40
	3,4,5	0.538	23.866	27
	4,5,	0.827	7.313	16
	5	0.997	0.116	7

Table A.4. (Continued)

	Roots	Wilks' λ_k	Test statistic[b] value (B_r)	D.F.[c]
07	1,2,3,4,5	0.010	177.299*	55
	2,3,4,5	0.114	83.605*	40
	3,4,5	0.352	40.199	27
	4,5,	0.841	6.667	16
	5	0.974	1.014	7
08	1,2,3,4,5	0.007	191.031*	55
	2,3,4,5	0.111	84.632*	40
	3,4,5	0.448	30.914	27
	4,5,	0.845	6.484	16
	5	0.993	0.270	7
11	1,2,3,4,5	0.014	164.345*	55
	2,3,4,5	0.138	76.249*	40
	3,4,5	0.479	28.338	27
	4,5,	0.745	11.333	16
	5	0.975	0.975	7
12	1,2,3,4,5	0.013	167.198*	55
	2,3,4,5	0.144	74.611*	40
	3,4,5	0.539	23.794	27
	4,5,	0.861	5.762	16
	5	0.994	0.232	7
13	1,2,3,4,5	0.034	130.184*	55
	2,3,4,5	0.313	44.720	40
	3,4,5	0.722	12.541	27
	4,5,	0.887	4.616	16
	5	0.979	0.817	7

b $B_r = -[n - 1 - 0.5(p + q + 1)] \ln \lambda_k$

c D.F. : Degrees of freedom = (p - k)(q - k).

* Significant at 5 % significant level.

Table A.5. Contribution of individual canonical correlations.

Tef Genotype	k	Contribution %	Cumulative %
01	1	74.30	74.30
	2	16.06	90.36
	3	7.50	97.86
	4	2.12	99.98
	5	0.02	100.00
03	1	70.92	70.92
	2	22.48	93.40
	3	4.54	97.95
	4	2.04	99.99
	5	0.01	100.00
04	1	78.56	78.56
	2	15.41	93.96
	3	3.58	97.54
	4	2.43	99.97
	5	0.03	100.00
05	1	84.83	84.83
	2	12.10	96.93
	3	1.89	98.83
	4	1.15	99.97
	5	0.03	100.00
06	1	79.50	79.50
	2	15.53	95.03
	3	3.58	98.61
	4	1.37	99.98
	5	0.02	100.00
07	1	73.34	73.34
	2	15.16	88.50
	3	10.15	98.65
	4	1.16	99.81
	5	0.19	100.00

Table A.5. (Continued)

Tef Genotype	k	Contribution %	Cumulative %
08	1	76.87	76.87
	2	17.13	94.00
	3	4.98	98.98
	4	0.98	99.96
	5	0.04	100.00
09	1	79.53	79.53
	2	14.56	94.09
	3	3.46	97.55
	4	2.41	99.96
	5	0.04	100.00
10	1	75.73	75.73
	2	16.99	92.71
	3	5.42	98.13
	4	1.26	99.39
	5	0.61	100.00
11	1	71.56	71.56
	2	20.93	92.49
	3	4.68	97.18
	4	2.61	99.78
	5	0.22	100.00
12	1	74.26	74.26
	2	20.17	94.43
	3	4.39	98.82
	4	1.14	99.96
	5	0.04	100.00
13	1	83.13	83.13
	2	13.28	96.41
	3	2.31	98.73
	4	1.06	99.79
	5	0.21	100.00

Table A.6. Canonical variate analysis of grain yield and component characters for 12 tef genotypes and environment variables.

Variable	Canonical variate CV_1	CV_2
Environment	Genotype 01	
Y_1	0.800	0.368
Y_2	-0.104	0.615
Y_3	-0.695	-0.479
Y_4	-0.784	-0.482
Y_5	-0.617	0.680
Y_6	-0.012	0.741
Y_7	-0.049	0.737
Y_8	-0.186	0.737
Y_9	-0.422	0.432
Y_{10}	0.701	0.434
Y_{11}	0.542	-0.323
Variance	0.284	0.251
Plant		
X_1	0.778	-0.024
X_2	0.687	0.382
X_3	-0.121	0.629
X_4	-0.706	0.308
X_5	-0.644	0.124
Redundancy	0.401	0.131
Environment	Genotype 03	
Y_1	0.775	0.344
Y_2	-0.381	0.548
Y_3	-0.665	-0.434
Y_4	-0.758	-0.450
Y_5	-0.671	0.648
Y_6	0.012	0.623
Y_7	-0.179	0.235
Y_8	-0.185	0.618
Y_9	-0.551	0.544

Table A.6. (Continued)

variable	Canonical variate	
	CV_1	CV_2
	Genotype 03	
Plant		
X_1	0.898	-0.021
X_2	0.717	0.417
X_3	-0.257	0.629
X_4	-0.622	0.114
X_5	-0.651	- 0.144
Redundancy	0.439	0.121
Environment	Genotype 04	
Y_1	0.753	0.331
Y_2	-0.341	0.576
Y_3	-0.636	-0.426
Y_4	-0.729	-0.430
Y_5	-0.694	0.606
Y_6	-0.076	0.602
Y_7	-0.055	0.188
Y_8	-0.272	0.596
Y_9	0.467	0.361
Y_{10}	-0.646	0.492
Y_{11}	0.576	0.379
Variance	0.287	0.198
Plant		
X_1	0.893	0.020
X_2	0.758	0.387
X_3	-0.304	0.387
X_4	-0.700	-0.311
X_5	-0.695	-0.126
Redundancy	0.488	0.083

Table A.6. (Continued)

Variable	Canonical variate	
	CV_1	CV_2
Environment	Genotype 05	
Y_1	0.854	0.159
Y_2	-0.323	0.859
Y_3	-0.765	-0.271
Y_4	-0.843	-0.285
Y_5	-0.562	0.781
Y_6	0.095	0.659
Y_7	-0.065	0.142
Y_8	-0.107	0.701
Y_9	-0.404	0.557
Y_{10}	-0.764	-0.228
Y_{11}	0.536	-0.478
Variance	0.319	0.215
Plant		
X_1	0.892	-0.123
X_2	0.899	0.187
X_3	-0.176	0.605
X_4	-0.646	0.440
X_5	-0.664	0.169
Redundancy	0.498	0.128
Environment	Genotype 06	
Y_1	0.856	0.216
Y_2	-0.375	0.726
Y_3	-0.766	-0.337
Y_4	-0.837	-0.335
Y_5	-0.561	0.812
Y_6	0.043	0.676
Y_7	0.057	0.235
Y_9	-0.311	0.543
Y_{10}	-0.739	-0.268
Y_{11}	0.450	-0.546
Variance	0.305	0.248

Table A.6. (Continued)

Variable	Canonical variate	
	CV_1	CV_2
	Genotype 06	
Plant		
X_1	0.839	-0.083
X_2	0.845	0.282
X_3	-0.147	0.608
X_4	-0.714	0.313
X_5	-0.765	0.066
Redundancy	0.507	0.133
Environment	Genotype 07	
Y_1	0.851	0.271
Y_2	-0.301	0.880
Y_3	-0.766	-0.384
Y_4	-0.848	-0.341
Y_5	-0.522	0.714
Y_6	0.118	0.637
Y_7	-0.048	0.225
Y_8	-0.080	0,651
Y_9	-0.367	0.576
Y_{10}	-0.762	-0.310
Y_{11}	0.502	-0.475
Variance	0.508	0.217
Plant		
X_1	0.821	-0.258
X_2	0.804	0.216
X_3	-0.175	0.270
X_4	-0.644	0.066
X_5	-0.686	0.069
Redundancy	0.447	0.099

Table A.6. (Continued)

Variable	Canonical variate CV_1	CV_2
Environment	Genotype 08	
Y_1	0.847	0.142
Y_2	-0.371	0.770
Y_3	-0.753	-0.244
Y_4	-0.835	-0.279
Y_5	-0.577	0.747
Y_6	0.085	0.625
Y_7	-0.077	0.135
Y_8	-0.117	0.666
Y_9	-0.421	0.532
Y_{10}	-0.752	-0.209
Y_{11}	0.551	-0.459
Variance	0.321	0.194
Plant		
X_1	0.904	-0.096
X_2	0.874	0.226
X_3	-0.096	0.739
X_4	-0.679	0.489
X_5	-0.746	0.057
Redundancy	0.522	0.170
Environment	Genotype 09	
Y_1	0.768	0.242
Y_2	-0.376	0.586
Y_3	-0.651	-0.310
Y_5	-0.721	0.578
Y_6	-0.053	0.528
Y_7	-0.134	0.186
Y_8	-0.254	0.535
Y_9	-0.447	-0.469
Y_{10}	-0.674	-0.272
Y_{11}	0.650	-0.382
Variance	0.316	0.155

Table A.6. (Continued)

Variable	Canonical variate CV_1	CV_2
Plant	Genotype 09	
X_1	0.802	0.125
X_2	0.692	0.449
X_3	-0.284	0.672
X_4	-0.751	0.318
X_5	-0.751	0.095
Redundancy	0.466	0.156
Environment	Genotype 10	
Y_1	0.879	0.114
Y_2	-0.357	0.833
Y_3	-0.805	-0.233
Y_4	-0.872	-0.248
Y_5	-0.458	0.801
Y_6	0.137	0.659
Y_7	0.064	0.118
Y_8	-0.062	0.711
Y_9	-0.246	0.551
Y_{10}	-0.775	-0.193
Y_{11}	0.395	-0.481
Variance	0.306	0.209
Plant		
X_1	0.416	0.143
X_2	0.723	0.293
X_3	0.001	0.668
X_4	-0.582	0.429
X_5	-0.613	0.212
Redundancy	0.282	0.156

Variable	Canonical variate	
	CV_1	CV_2
Environment	Genotype 11	
Y_1	0.759	0.518
Y_2	-0.414	0.753
Y_3	-0.638	-0.608
Y_4	-0.711	-0.606
Y_5	-0.761	0.575
Y_6	-0.085	0.650
Y_7	-0.149	0.279
Y_8	-0.287	0.607
Y_9	-0.583	0.533
Y_{10}	0.251	-0.431
Y_{11}	-0.663	-0.518
Variance	0.290	0.268
Plant		
X_1	0.798	0.011
X_2	0.571	0.461
X_3	-0.331	0.598
X_4	-0.717	0.134
X_5	-0.647	-0.193
Redundancy	0.401	0.125
Environment	Genotype 12	
Y_1	0.798	0.205
Y_2	-0.387	0.567
Y_3	-0.695	-0.310
Y_4	-0.791	-0.333
Y_5	-0.597	0.757
Y_7	-0.063	0.253
Y_8	-0.163	0.648
Y_9	-0.421	0.625
Y_{10}	-0.705	-0.246
Y_{11}	0.541	-0.533

Table A.6. (Continued)

Variable	Canonical variate CV_1	CV_2
Variance	0.295	0.261
	Genotype 12	
Plant		
X_1	0.844	-0.134
X_2	0.786	0.306
X_3	-0.262	0.735
X_4	-0.660	0.349
X_5	-0.736	-0.042
Redundancy	0.475	0.155
	Genotype 13	
Environment		
Y_1	0.843	0.290
Y_2	-0.407	0.765
Y_3	-0.748	-0.413
Y_4	-0.830	- 0.417
Y_5	-0.569	0.768
Y_6	0.076	0.768
Y_7	-0.048	0.017
Y_8	-0.125	0.819
Y_9	-0.394	0.425
Y_{10}	-0.749	-0.388
Y_{11}	0.525	-0.330
Variance	0.315	0.252
Plant		
X_1	0.762	0.063
X_2	0.796	0.258
X_3	-0.274	0.633
X_4	-0.589	0.314
X_5	-0.667	0.081
Redundancy	0.417	0.115

Table A.7. Principal components analysis of yield and component characters for 12 tef genotypes grown under 12 environments.

Genotype	Parameter	Loading	Principal Component				
			PC_1	PC_2	PC_3	PC_4	PC_5
01	Eigenvalue		2.52	1.57	0.49	0.25	0.02
	% variation explained		50.42	31.41	9.83	4.91	3.43
		X1	-0.70	0.58	0.31	0.16	0.22
		X2	-0.59	0.75	0.01	-0.19	-0.25
		X3	0.49	0.71	-0.49	0.13	0.08
		X4	0.85	0.32	0.21	-0.32	0.15
		X5	0.85	0.26	0.33	0.25	-0.18
03	Eigenvalue		2.61	1.40	0.63	0.19	0.17
	% variation explained		52.11	27.97	12.68	3.76	3.48
		X1	-0.81	0.43	0.27	0.13	0.27
		X2	-0.66	0.67	0.17	-0.14	-0.25
		X3	0.40	0.74	-0.42	0.15	0.01
		X4	0.83	0.43	0.17	-0.28	0.16
		X5	0.82	0.18	0.48	0.23	-0.11
04	Eigenvalue		2.88	1.17	0.43	0.30	0.22
	% variation explained		57.6	23.4	8.7	5.9	4.4
		X1	-0.80	0.44	0.22	-0.09	0.32
		X2	-0.75	0.54	0.15	0.18	-0.29
		X3	0.52	0.76	-0.34	-0.21	-0.02
		X4	0.84	0.31	0.07	0.41	0.14
		X5	0.83	0.13	0.49	-0.22	-0.09

Table A.7. (Continued)

Genotype	Parameter	Loading	Principal Component				
			PC_1	PC_2	PC_3	PC_4	PC_5
06	Eigenvalue		2.84	1.44	0.46	0.17	0.09
	% variation explained		56.9	28.7	9.2	3.3	1.9
		X1	-0.78	0.48	0.32	-0.23	-0.09
		X2	-0.74	0.59	0.09	0.27	0.09
		X3	0.37	0.82	-0.42	-0.12	0.04
		X4	0.87	0.38	0.16	0.14	-0.20
		X5	0.88	0.20	0.38	-0.07	0.18
07	Eigenvalue		2.63	1.52	0.54	0.17	0.14
	% variation explained		52.6	30.4	10.9	3.3	2.8
		X1	-0.73	0.59	0.17	0.29	0.01
		X2	-0.62	0.72	0.13	-0.28	-0.02
		X3	0.49	0.65	-0.46	0.03	0.07
		X4	0.84	0.36	0.17	0.03	-0.28
		X5	0.87	0.28	0.39	-0.01	0.23
08	Eigenvalue		3.02	1.25	0.37	0.22	0.14
	% variation explained		60.4	25.1	7.4	4.3	2.8
		X1	-0.85	0.30	0.37	-0.08	-0.19
		X2	-0.77	0.56	0.05	0.18	0.23
		X3	0.40	0.86	-0.24	-0.20	-0.05
		X4	0.87	0.32	0.11	0.33	-0.12
		X5	0.89	0.07	0.40	-0.16	0.16

Table A.7. (Continued)

Genotype	Parameter	Loading	Principal Component				
			PC_1	PC_2	PC_3	PC_4	PC_5
10	Eigenvalue		2.08	1.55	0.81	0.42	0.13
	% variation explained		41.5	31.1	16.3	8.4	2.7
		X1	-0.19	0.72	0.59	-0.29	-0.03
		X2	-0.16	0.87	-0.14	0.44	0.04
		X3	0.55	0.51	-0.56	-0.34	0.05
		X4	0.95	0.07	0.11	0.13	-0.26
		X5	0.90	-0.08	0.33	0.08	0.24
11	Eigenvalue		2.36	1.74	0.59	0.21	0.10
	% variation explained		47.3	34.7	11.8	4.2	2.0
		X1	-0.67	0.63	0.29	0.25	0.12
		X2	-0.39	0.87	0.29	-0.26	-0.13
		X3	0.48	0.66	0.08	0.17	-0.02
		X4	0.89	0.35	0.14	-0.17	0.20
		X5	0.87	0.15	0.42	0.15	-0.16
12	Eigenvalue		2.74	1.32	0.53	0.25	0.16
	% variation explained		54.8	26.4	10.6	5.1	3.1
		X1	-0.83	0.73	0.17	-0.26	0.03
		X2	-0.60	0.75	0.03	0.27	0.00
		X3	0.54	0.52	-0.66	-0.04	0.05
		X4	0.83	0.45	0.22	-0.01	-0.25
		X5	0.86	0.28	0.34	0.04	0.24

Table A.7. (Continued)

Genotype	Parameter	Loading	Principal Component				
			PC_1	PC_2	PC_3	PC_4	PC_5
13	Eigenvalue		2.45	1.65	0.63	0.14	0.12
	% variation explained		49.0	33.0	12.6	2.9	2.5
		X1	-0.60	0.73	0.17	-0.26	0.03
		X2	-0.60	0.75	0.03	0.27	0.00
		X3	0.54	0.52	-0.66	-0.04	0.05
		X4	0.83	0.45	0.22	-0.01	-0.25
		X5	0.86	0.28	0.34	0.04	0.24