A COMPARISON OF PATH-COEFFICIENT ANALYSIS AND PRINCIPAL COMPONENT ANALYSIS IN TURMERIC

BY

GEETHA, U.

THESIS
Submitted in partial fulfilment of the requirements
for the degree
Master of Science (Agricultural Statistics)
Faculty of Agriculture
Kerala Agricultural University

Department of Statistics
COLLEGE OF VETERINARY AND ANIMAL SCIENCES
Mannuthy, Trichur
1985

DECLARATION

I hereby declare that thesis entitled "A COMPARISON OF PATH-COEFFICIENT ANALYSIS AND PRINCIPAL COMPONENT ANALYSIS IN TURMERIC " is a bonafide record of research work done by me during the course of research and that thesis has not previously formed the basis for the award to me of any degree, diploma, associateship, fellowship, or other similar title, of any other University or Society.

Geetha.u
GEETHA U

MANNUTHY
16- 10- 1985.

CERTIFICATE

Certified that this thesis entitled "A COMPARISON OF PATH-COEFFICIENT ANALYSIS AND PRINCIPAL COMPONENT ANALYSIS IN TURMERIC" is a record of research work done independently by MISS. U. GEETHA under my guidance and supervision and that it has not previously formed the basis for the award of any degree, fellowship, or associateship to her.

(P.V.PRABHAKARAN)
Chairman
Advisory Committee
Professor of Statistics,
College of Agriculture,
Vellayani.

Vellayani,
16-10-1985.

APPROVED BY

Chairman:

Shri. P.V.PRABHAKARAN

Members:

Shri.V.K.GOPINATHAN UNNITHAN

Shri. E.V.NYBE

Dr.K.M.NARAYANAN NAMBOOTHIRI

External Examiner:

22/2/86

ACKNOWLEDGEMENT

I express my deep sense of gratitude and indebtedness to Shri P.V.Prabhakaran, Professor of Statistics, College of Agriculture, Vellayani, for suggesting the research problem, his expert guidance, constant encouragement and meticulous efforts rendered during the entire course of the research work and preparation of thesis.

I extend my sincere gratitude to Shri V.K. Gopinathan Unnithan, Associate Professor of Statistics, Shri Joseph Philip, Assistant Professor of Plantation crops and Dr.K.M.Narayanan Namboothiri, Professor of Agricultural Botany, College of Horticulture, Vellanikkara, the members of the advisory committee, who took keen interest in going through the material and making useful suggestions.

I wish to express my heartfelt thanks to the Head of the Department of Plantation Crops, College of Horticulture, Vellanikkara, for providing the necessary data for the study.

I am thankful to the staff and students of the Department of Statistics, friends and all others for their timely help and encouragement to the successful completion of this thesis.

Mr.L.Radhakrishnan Potti, Senior Office Superintendent also deserves my thanks for neatly typing the manuscript.

I also thank the Kerala Agricultural University for the award of junior fellowship during the period of the Post-graduate programme.

GEETHA, U.

CONTENTS

--:0:--

INTRODUCTION

1. INTRODUCTION

Statistical methods as applied to the field of crop and animal sciences are mainly intended to increase production by developing and implementing suitable techniques for the efficient utilisation of the available resources. They help the breeder of plants and animals to get a clear understanding of the absolute criteria on the basis of which desirable genotypes could be evolved. Production of food is the major concern through out the world. To meet with the food requirements of the coming generation, it is necessary to enhance crop and livestock production. Thus it is needless to say that any crop improvement programme should have as its prime objective the testing and evolution of new genotypes superior to the existing ones in production and consumption. Selection, the most common method of crop improvement is an integral phase of any breeding programme and has been applied to agricultural crops for hundreds of years.

A suitable selection criterion could be evolved only if the researcher possesses a sound knowledge of the various genetic parameters involved, the economic goals to be achieved and the appropriate measures the genotypic and phenotypic performances of the cultivars. Any breeding programme involves two steps (1) choice of a suitable mating system (2) selection of individuals to be used as parents. The choice of the mating plan depends on the genetic variation present in the material that is to be improved. If the variation is mostly additive selection of parents with random mating is preferred. Improvement brought about by selection not only increases the values of the genotypes for the desired traits but also enhances the usefulness of the plant as potential parents for hybridisation. In order to make the process of selection effective, it would be necessary that members of the population on which selection is to be practiced should vary in their genetic make up with regard to the desired characters. A permanent improvement of the desired characters could be achieved only through exploiting the genetically deterministic variation. Thus a basic pre-requisite to the planning of any breeding programme is to determine how much proportion of the total variability in character would be caused by the differences in the genetic make up of the individuals of the crop species. A quantitative measure of this is provided by the coefficient of heritability.

Heritability is of interest to plant breeder not only as a measure of the value of selection for the specific character in various types of progenies but also an index of transmissibility of the trait to the very next generation. The fraction of total variability which is heritable due to genetic make up is termed as the coefficient of heritability in the broad sense (Lush, 1937) where as the proportion of the variability that is due to additive genetic variance is called the coefficient of heritability in the narrow sense. If the genotype completely determines the phenotype both the variance would be equal and heritability would be 100 percentage. A knowledge of the magnitude of heritability gives an idea about the genetic improvement which could be brought about through selection. Heritability also indicates the accuracy with which selection for the genotype could be made from a phenotype of the individual or a group of individuals.

Another apriori basic information needed is about the extent of association between genetic values of a pair of quantitative traits. Generally estimates of genetic correlations are of very low precision. A knowledge of the magnitude and sign helps to understand how the improvement in one character will cause simultaneous change in other characters. A comparison of phenotypic and genotypic correlations would give an indication of the effect of environment on the genetic performance of the individuals of the population.

The third genetic parameter which needs consideration is the selection differential which indicates the vigour with which a breeder practices selection. When selection differential is expressed in terms of phenotypic standard deviation, it is known as the intensity of selection. The genetic improvement per generation or genetic advance is another parameter which is simply the product of heritability and selection differential. It is an indirect measure of the expected change in the population that can be brought about by selection on the particular trait.

The principles of individual plant selection to establish varieties of self pollinated crops were first developed at the plant breeding station of the Swedish Seed Association, Svalof, shortly before 1900. Selection on the basis of individuality means the individuals are selected for breeding purposes on the basis of their own phenotypic performances. The principle of individual selection is also applied in mass selection, since the selected individuals are used enmasse for mating. In mass selection individuals are selected solely in accordance with their phenotypic values. Another method of selection is family selection which consists of selecting a group of promising families as units, in accordance with their mean phenotypic values. This can be practiced when the population can be divided into families of full sibs or half sibs, each family providing a mean phenotypic value. Sib selection is selection of individuals on mean values of sibs. Pedigree selection is another method which is based on the mean performance of the two parents or on the mean performance of the parents and grand parents. Progeny selection is the selection of individual on the basis of the performance of the new generation. But all these methods of selection are useful for the identification of a superior plant or genotype based on single character alone.

The breeder is seldom faced with a situation in which modification of only a single attribute is desired. Normally modification will be desired on several traits which may or may not be interrelated. It is therefore necessary that selection is applied simultaneously to all traits of economic importance. Tandem selection, independent culling levels and selection index are the three commonly used methods of simultaneous selection.

Tandem method of selection is practiced only for one trait at a time until satisfactory improvement was achieved in that trait. Then selection efforts could be directed towards the second trait. In the second method of selection viz. independent culling levels selection is made simultaneously for all the characters but independently, rejecting all individuals that fail to meet the minimum standard for any one trait. Hazel and Lush (1942) have compared the three methods of selection and concluded that the method of selection index is more efficient than the other two methods. In this method the net merit of an individual is defined on the basis of an index by combining scores for each component character. In situations when more than one basis of selection is applied, for choosing effective criteria, it is necessary to compare their relative efficiencies. The choice would obviously fall on one which result in maximum genetic gain per generation.

Statistical methods play a vital role in evolving an appropriate selection criterion - a selection index for the identification of superior genotypes. The methods of multivariate analysis are universely used in this direction. An important problem in multivariate analysis is to identify the major variables to be retained in the linear selection model. When observations are collected on several variables, they may be interrelated. A series of unvariate analysis made on each character separately ignores the possible correlations among sets of variables and lead to misinterpretation of the data. Multiple regression analysis makes use of the assumption that the explanatory variables of the model are not strongly correlated. 'Multicollinearity' may lead to biased and often difficult to interpret estimates of partial regression coefficients and their errors. In order to remove multicollinearity and to reduce the large number of correlated quantitative traits into relatively smaller set of statistically independent linear combinations, the method of principal component analysis is usually attempted. A simple variant form of principal component analysis is factor analysis. In fact, factor analysis performed using the principal axis method is essentially the same as the principal component analysis. Principal component analysis is used to get a better configuration of the multivariate structure. The causative sources of common variation can be identified. Weights may be assigned to each variable so that resultant compositive variable as a set may have maximum variance. These derived composite variables will be mutually orthogonal. As factor analysis ignores dependence relationship, various traits can be evaluated perse rather than in terms of their relationship with the dependent character.

In practice, principal component analysis is one way of reducing variables that are to be retained in the model. Regression analysis is based on previously assumed relationship where as principal component and factor analysis are useful in identifying unknown relationships between variables. In factor analysis and principal component analysis normality is not necessary. Further in principal component analysis and factor analysis the cause effect relationship of structural variables is not assumed.

Another approach to the same problem is through the method of path-coefficient analysis. Path analysis consists of a series of multiple regression analyses with an additional assumption of causal relationship between the independent and dependent variables. In path analysis an over all measure of the emperical relationship is decomposed into a series of paths. In fact, path-coefficients are standardised regression coefficients and are used to measure the direct and indirect influences of one variable upon another in the study of specific forces acting to produce a given correlation. All the assumption which are implied in the method of multiple regression are inherent in path-coefficient analysis also. This method deals with a closed system of variables that are immediately linearly related. The separation of correlation coefficient into various components is one of the accomplishments of the method of path-coefficients. Analogus to the analysis of variance, the path method may be called the analysis of correlation.

Path-coefficient analysis, though simple, highly informative and popular among breeders, fail to produce meaningfull results in certain situations especially when confronted with extreme multicollinearity. Further path-coefficient analysis is based on strict assumptions and the validity of the result becomes dubious in cases where such assumptions are not satisfied.

Principal component analysis is a more general method and is extensively flexible in dealing with multivariate data. In this study a comparison is made using the experimental data of a field trial on the relative merits and limitations of these two accepted techniques of multivariable analysis.

Curcuma longa L. is the important species of turmeric grown in India which belongs to the natural order Scitaminae and family Zingiberaceae. In India, it is mainly valued for use as spices and also used in medicines. Among the spices, turmeric ranks fourth with regard to the foreign exchange earnings next to black pepper, cardamom and ginger. India is the largest producer and exporter of turmeric. In India, turmeric is cultivated in an area of 88200 hectares and of this 4000 hectares (4.5 percentage) is in Kerala State (Anon., 1983). Andra Pradesh, Tamil Nadu, Bihar and Orissa are the turmeric producing states other than Kerala. Total annual production of turmeric in India is about 172600 tonnes, Kerala State contributes about 4.4 percentage of total production of turmeric and add 76 lakhs of rupees annually (Anon., 1983).

In spite of its importance as a commercial crop and a foreign exchange earner, present yield rate of turmeric in India is pathetically low as compared to that in other countries. The cost of cultivation of turmeric has been increased due to ever enhanced labour cost during recent years. It is a shade tolerant crop and there are possibilities of growing it as an intercrop in coconut gardens. Both morphological and physiological factors play an important and inter dependent role in determining the yield rate of the crop. The best yields obtainable with the present cultivars of turmeric are relatively low when compared to most of the other field crops. In order to bring about marked improvement in yield and curcumin content, the breeder has to identify desirable genotypes for breeding purposes. The present study based on the data collected from the Department of Plantation Crops, College of Horticulture, Vellanikkara pertaining to a field experiment on turmeric is therefore undertaken with following objectives:

1. To summarise most of the variation in a multivariate system in a fewer number of uncorrelated variables.

2. To measure the relative contribution of each individual character affecting the yield of turmeric.
3. To construct selection indices with different system of weighting and compare their relative efficiencies.
4. To evolve suitable selection criteria for varietal screening and evaluation.
5. To compare the relative advantages and limitations of two widely accepted techniques of multivariate analysis with reference to actual field data.

REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

Even though very little work on the breeding and genetics of turmeric has been attempted in India or elsewhere, numerous reports on such investigation on other crops are available and a brief review of literature of studies conducted on these aspects on turmeric and also on other crops is furnished below.

2.1. GENETIC PARAMETERS AND THEIR ESTIMATION

Galton (1889) conceived the idea of correlation of variables for the first instance.

Fisher (1918, 1954) developed the method of applying the theory of correlation of variables, in the understanding of their influence on biological systems.

Lush (1949) devised an accurate and easily manageable method for the calculation of the genetic advance under specified intensities of selection.

Robinson *et al.* (1951) computed all possible genotypic and phenotypic correlation coefficients from a combined analysis of different populations of prolific corn and among the study variables ears per plant had showed the highest positive genotypic correlation with yield.

Burton (1952) introduced a simple and convenient procedure for the calculation of the phenotypic and genotypic coefficients of correlation utilising data from breeding trials.

Warner (1952) introduced a new method of estimating heritability from the variances of segregating populations.

Johnson *et al*. (1955) introduced the methodology for partitioning the total variance into that due to genotype, phenotype and error in the analysis of variance.

Purewal (1957) obtained highly significant positive correlation between plant height and yield in tuberous plants like colocasia.

Al-jibouri *et al*. (1958) estimated the genotypic and environmental variances and covariances in upland cotton and genotypic correlation indicated that a positive relationship existed between lint yield and lint percentage.

Robertson (1959) derived formulae for the sampling variances of the genetic correlation coefficient. The formulae were very similar to those for the sampling variance of heritability from the same experiment except that the denominator contained intraclass correlation coefficients.

Jones (1959) reported that yield of tuber in tapioca was associated positively with relatively larger leaves, longer duration and higher values for the leaf area index.

Gill and Jensen (1968) derived the probability of obtaining negative estimates of heritability in full sib and half sib analysis for balanced nested models.

Meister and Thompson (1976) studied phenotypic correlation among yield and its components in potato.

Maity and Chatterjee (1977) found that yield of potato was positively correlated with height and number of tubers per plant.

Kamalam *et al.* (1977) conducted a study in ten varieties of sweet potato and estimated the genetic parameters. Length of vine and number of tubers per plant showed very high degree of phenotypic and genotypic coefficient of variation. Number of tubers had positive significant correlation on yield.

Lin and Alliare (1977) defined the heritability of an index as the regression of the genetic index on the selection index. The genetic index was obtained by substituting the genetic values of the index traits into the phenotypic values.

Nigam et al. (1979) proposed an estimator of heritability obeying an apriori restriction on the range of its value.

Mohanty (1979) studied correlation coefficients among different characters and yield components in turmeric and revealed that tall plants with more number of broad leaves likely to produce high yielding turmeric types.

Genetic variability and heritability were estimated by Mohanty and Sarma (1979) for a number of characters in different cultivars of ginger. Their study indicated that straight selection can be made to improve almost all characters except the number of tertiary fingers and 'straw' yield.

Nambiar (1979) estimated the intercorrelations among the morphological characters and yield in turmeric and the results showed that number of tillers, plant height and number of fingers had high significant positive correlations with the yield of turmeric. Further, the morphological characters showed high significant intercorrelations among themselves.

From an overall analysis of path-coefficients, Usharani and Rao (1981) suggested the characters as major contributors towards yield as those which had positive direct effects and those having small negative direct effect but high genotypic correlation with yield.

Boomikumaran and Rathinam (1981) emphasised on those characters which had positive direct effects and those having small negative indirect effect but high genotypic correlation with yield in building up a high yield strain in green gram.

Joseph Phillip and Nair (1983) studied the relationship between yield and other morphological and rhizome characters and revealed that height of the plant petiole length, length and breadth of leaf, number of leaves per tiller, length of primary fingers and girth of mother rhizome at centre were positively correlated with the yield.

2.2. SELECTION OF CHARACTERS

2.2.1. PRINCIPAL COMPONENT ANALYSIS AND FACTOR ANALYSIS

Thurstone (1931) described a more generally applicable method of factor analysis which had no restrictions as regards to group factors and which did not restrict the number of general factors that are operative in producing the intercorrelations.

According to Hotelling (1933) Principal components are linear combinations of statistical variables which have special properties in terms of variance. The first principal component is the normalised linear combination with maximum variance.

Girshick (1936) has shown that principal components are linear functions of variables which have least variance ascribable to errors of measurement and factor loadings of principal components are maximum likelihood statistics.

Rao (1955) explained the experimental situation and nature of the data on which the technique of factor analysis can be successfully employed.

The normalised raw varimax criterion for the rotation of principal components was proposed by Kaiser (1956, 1958). He further showed that the varimax rotation satisfied the desirable property of factorial invariance.

Kaiser (1960) has stated that a whole of host of criteria involving both statistical and practical considerations suggest the number of principal components having associated latent roots greater than Si^2/n(=1 when principal component analysis is performed on the correlation matrix) as the best single criterion for the number of factors to be assumed in the description of any structure whether principal component analysis or Factor analysis have been performed.

Lawley (1963) developed a statistical test for testing the equality of the last (m-1) characteristics roots of the correlation based principal component analysis.

Lawley and Maxwell (1963) developed the least square methods for rotating the factor structures.

According to Cattel (1965) utility of the factor analysis lies not only of explanatory stages of research but also at later stages when simultaneous action of several factors influencing a variable is to be critically analysed.

Pearce (1967) has described an interesting application of principal component analysis in apple for the prediction of the tree weight on the basis of relevant size measurements of individual tree.

The pattern of diversity in the genus sorghum was analysed using centroid method of factor analysis based on ten to twelve groups of populations by Murthy and Arunachalam (1967). The factor loadings on the variables were different in two groups of populations. Three factors were found to be adequate to account for more of intercorrelation in both the genotypic and environmental correlation matrices.

Walton (1972) studied the relative variation of characters in a five by five diallel cross among varieties of spring wheat using factor analysis and found that factors concerned with flag leaf area and duration were the most important contributers towards the total divergence.

Principal component analysis was performed on twenty two morphological characters and yield determining characters of sixteen cultivars and strains of dry beans by Denis and Adams (1978). Variables with the highest positive loadings on the first axis were seed weight, pod thickness, pod breadth, pod length and internode length.

Tikka and Asawa (1978) used correlations of twenty eight genotypes for factor analysis and only two factors were found important in explaining the variations in the seven traits.

Johnston (1978) has summarised the advantage of principal component analysis as (1) to identify groups of uncorrelated variables (2) to reduce number of variables (3) to remove multicollinearity.

Agarwal <u>et al.</u>(1980) used the technique of Principal components to develop weather indices. Further they used principal components as independent variables in a multiple regression equation. The first two principal components accounted for about 80 percentage of total variation in yield.

Mahajan <u>et al.</u>(1981) applied the analysis through the principal component technique and showed that the most of the variations in late duration cultures of rice could be explained by ear-bearing tillers or grain number per panicle and 100 grain weight among a set of six variables.

Factor analysis was carried out by Tikka and Asawa (1982) in cowpea and showed that the complex relationship among the different traits could be explained with the help of three factors.

2.2.2. PATH-COEFFICIENT ANALYSIS

The theory of path-coefficient analysis was first established by Wright (1921).

Li (1956) has pointed out that the most direct applications of path-coefficient method is the deduction of correlation between two variables which are linear functions of some common variables.

Kempthorne (1957) has illustrated three main applications of path-coefficients. They were (1) To study the consequences of a number of linearly related forces, on their resultant in a cause and effect system (2) To examine the feasibility of pattern of causal forces in estimating path-coefficients whose direct results cannot be measured (3) In making clear, what can be deducted from a set of correlation coefficients.

Dewey and Lu (1959) used path-coefficient analysis to study the relative importance of six biometric characters in enhancing wheat grass production. Fertility and plant size had strong direct and indirect influences on yield.

Wright (1968) suggested suitable transformations in cases where the underlying relationship between variables is non-linear.

Butany <u>et</u> <u>al</u>.(1968) used path analysis to identify yield components in cotton and revealed that sympodia and boll weight contributed most direct and indirect influences on yield.

Badwal <u>et</u> <u>al</u>. (1970) revealed by path analysis that capsule number and 1000-seed weight were the major factors which directly contributed to seed yield in linseed.

Norman (1972) used path-coefficient analysis to identify important components of sugarcane production and found that number of millable stalk per unit area was the most important factor followed by stalk diameter and stalk length.

Khangura and Sandhu (1972) used path analysis in groundnut and found that length of primary branches was the most important character having highest direct and indirect effects on yield.

Path analysis was attempted in flue-cured tobacco by Venkata Rao <u>et</u> <u>al</u>.(1973) and they found that plants with greater number of broad leaves would be ideal for selection.

Shivahare et al. (1975) used path analysis in Indian mustard and found that number of secondary and primary branches, number of seeds per siliqua and 1000 seed weight were the major components of seed yield.

Kamalam et al. (1977) conducted path coefficient analysis in sweet potato. Among the characters, number of tubers showed the maximum positive direct as well as indirect effects on yield of tubers.

Ratnambal (1979) applied path-coefficient analysis in ginger and revealed that the character 'plant height' exhibited a high direct effect as well as indirect effects in the establishment of correlations between yield and other morphological characters.

Nambiar (1979) applied path analysis to identify the important yield components of turmeric and among the characters, plant height and number of fingers were the major components of rhizome yield. He concluded that plant height in turmeric is a single important morphological character for which selection for yield could be made.

On the basis of correlation and path coefficient analysis Bhagchandini and Choudhury (1980) reported that diameter of the root is the most reliable selection cirterion for breeding better varieties in tropical carrot.

George (1981) applied path-coefficient analysis on open pollinated progenies of turmeric and found that length of secondary finger was the major factor contributing to yield of rhizome followed by number of primary fingers.

Through path analysis on groundnut, Nagabhushanam et al. (1982) found that number of mature pods and 100 kernel weight were the major contributors of yield while internodal length and number of primary branches exhibited only negative direct effects.

2.4. SELECTION INDEX

2.4.1. PHENOTYPIC INDEX

If information on genetic parameters are not estimable selection indices based on auxillary traits are often found to be more reliable than direct selection.

Bartlett (1933) developed a technique of multivariate regression analysis based on the linear model

$Y = b_0 + b_1 x_1 + \ldots\ldots\ldots\ldots\ldots + b_p x_p$
where the b's are so selected as to maximize the correlation between the obtained and expected values of the dependent variable.

The technique of multiple linear regression was originally proposed by Fisher (1936) to construct discriminant function of the form $Z = \sum biXi$.

Gupta (1971) used the multiple regression analysis to construct a reliable selection index for improving fodder quality in pearl millet and revealed that number of leaves, stem thickness, plant height and earliness accounted for 87 percentage variation in green fodder yield.

Narain and Mishra (1971) developed a new index called the 'Phenotypic index' which reduces error variation of main trait by using partial regression of several auxiliary traits.

To predict the yield of wheat Singh et al. (1976) fitted multiple regression models and calculated the co-efficient of determination and tested for significance.

Sandhu and Korla (1976) fitted a multiple regression equation to predict the yield of onion using the characters number of seeds/plant stalks per plant and weight per umbel.

Asawa et al. (1977) constructed different phenotypic indices in sunflower and found that index with the characters height, days to heading, head diameter and percentage filling in seeds explained maximum percentage of total variation.

Ratnambal (1979) applied multiple regression analysis using morphological characters in in ginger and revealed that the final yield could be fairly accurately estimated, based on certain morphological characters at 90th and 120th day after planting.

2.4.2. GENOTYPIC INDEX

The first application of the selection index to plant breeding was by Smith (1936). According to him genetic worth of an individual 'H' is defined as $H = \sum aiGi$ where Gi's are the genetic values and ai's are the weights depending on the economic values of corresponding characters. Another function 'I' is defined as $I = \sum biPi$ where bi's are the regression coefficients estimated such that the correlation between H and I is maximum.

The first application of the selection index to animal breeding was by Hazel (1943).

Kempthorne and Nordskog (1959) developed selection indices by imposing certain restrictions on the measured traits.

Brim et al. (1959) constructed selection indices by giving different sets of weights for selecting promising genotypes in Soyabean.

Williams (1962) constructed selection indices for selecting the best 'm' out of a random sample of 'k' linear functions of non-observable multivariate normal variables.

Harris (1964) developed a method to measure maximum attainable progress from index selection involving estimates of population parameters.

James (1968) gave a different approach to restricted selection. According to him it might be convenient not to measure certain traits whose inclusion added little to efficiency, even if they were of some economic importance.

Pesek and Baker (1969) suggested a method of constructing selection indices based on the principles of specifying ' desired gains' of a group of quantitative traits. Their method avoids the problem of assigning relative economic weights to the traits which some times limits the practical use of index selection.

Verma and Singh (1971) worked out selection indices to bring about improvement in barnyard millet. The selection index based on yield, number of tillers, number of nodes and days to flowering was the most efficient.

Pundir and Rai (1971) applied multiple selection criteria to bring about genetic improvement of toria populations. Selection index based on seeds per siliqua, 1000 seed weight and yield per plant was observed to be the most effective one.

The approach of discriminant function was used to construct selection indices for yield in Indian mustard by Singh and Singh (1974). Selection based on number of primary branches, siliqual length and plant height recorded the highest relative efficiency.

Selection indices for fodder yield in sorghum were computed for different combination of characters by Singh and Singh (1974). The best selection index, which included the leaf length, stem girth, stem length, number of inter-nodes and fodder yield resulted in 20-34 percentage increase of efficiency over direct selection.

Singh and Singh (1976) developed a set of discriminant functions in different generations of chilli for identification of superior genotypes. Days to flower, fruit length and number of fruits per plant were the major components in determining yield in chilli.

Chandrappa et al. (1978) constructed selection indices using discriminant function in triticale. Maximum efficiency was found for the index with 1000 kernel length, days to heading, yield, floral fertility and kernel length.

Selection indices for yield were constructed and their efficiencies assessed in terms of the expected genetic advance using 30 varieties in niger by Saha and Patnaik (1980). Five groups of selection indices were evaluated, the characters being choosen on five different criteria (1) heritability (2) phenotypic correlation with yield (3) genotypic correlation with yield (4) product of heritability and phenotypic correlation with yield (5) product of heritability and genotypic correlation with yield. The absolute value of the product of heritability and genotypic correlation with yield was found to be the most efficient criterion for the selection of characters.

MATERIALS AND METHODS

3. MATERIALS AND METHODS

The data for the study were gathered from a field trial on turmeric conducted by the Department of Plantation Crops, College of Horticulture, Vellanikkara, Trichur District during the year 1977-78. The experiment was laid out in randomised block design with four replications and 19 varieties as treatments. The weather and seasonal conditions during the period of study were more or less normal. The soil was red laterite loam. Adequate drainage was provided. Small elevated beds of height 0.25 cm and area 3.6 m x 1.6 m were raised providing channels of width 0.6 m around each to prevent the soil erosion and water logging. Thus the experimental plot consisted of a net area of 3 m x 1 m with a population of 48 plants at a uniform spacing of 25 cm x 25 cm.

The nineteen varieties of turmeric used for the study were the following.

1. Chayapasupa
2. Kuchupudi
3. Kodur
4. G.L.puram - II
5. Mannuthy local
6. Nandyal

7. Amruthpani kotapeta
8. Armoor
9. Duggirala
10. Amalapuram
11. Vontimitta
12. Kasturi Tanuka
13. Amruthapanikothapeta CII-317.
14. Dindrigam Ca-69
15. Duggirala CII-325
16. Rajapuri
17. Armoor CII-324
18. G.L.Puram-1
19. Tekurpeta

Observations on morphological and rhizome characters were collected from a random sample of 40 plants selected at the rate of ten per plot. The characters included in the study were the following.

1. Length of secondary finger.
2. Girth of secondary finger at centre.
3. Number of nodes per secondary finger.
4. Internodal distance of secondary finger.
5. Length of mother rhizome
6. Number of nodes per mother rhizome

7. Girth of mother rhizome at centre.
8. Internodal distance of mother rhizome
9. Number of nodes per primary finger.
10. Length of primary finger.
11. Internodal distance of primary finger.
12. Girth of primary finger at centre.
13. Length of last fully opened leaf.
14. Breadth of leaf at centre.
15. Petiole length.
16. Height of the plant.
17. Number of leaves per tiller.
18. Yield of raw rhizome per bed.

3.1. ESTIMATION OF GENETIC PARAMETERS

The progress of selection in a population is primarily conditioned by magnitude, nature, and interaction of genotypic and non-genotypic variations. This indicates the need of partitioning the total variability into heritable and non-heritable components so as to enable the breeder to adopt suitable breeding technology.

A skeleton of the Analysis of Variance table of the experiment is given below.

Table 3.1. Analysis of variance of the field trial on turmeric in R.B.D. with 'r' replication and 't' varieties.

Source	Degree of freedom.	Sum of squares	Mean sum of squares (M.S.S.)	Expected M.S.S.
Replications	r-1	S.S.R.	M.S.S.R.	
Varieties	t-1	S.S.T	M.S.S.T	$\sigma_e^2 + r\,\sigma_g^2$
Error	(r-1) (t-1)	S.S.E.	M.S.S.E	σ_e^2
Total	rt-1			

Phenotypic variance (σ_p^2), genotypic variance (σ_g^2) and environmental variance (σ_e^2) are estimated from table 3.1 as follows.

$$\hat{\sigma}_e^2 = \text{M.S.S.E} \quad (3.1)$$

$$\hat{\sigma}_g^2 = \frac{\text{M.S.S.T.} - \text{M.S.S.E.}}{r} \quad (3.2)$$

$$\hat{\sigma}_p^2 = \hat{\sigma}_g^2 + \hat{\sigma}_e^2 \quad (3.3)$$

These components of variance are further used for estimating other useful statistical measures such as phenotypic coefficient of variation, genotypic coefficient of variation, heritability, phenotypic and genotypic coefficients of correlation, genetic advance and genetic gain.

3.1.1. PHENOTYPIC COEFFICIENT OF VARIATION

Phenotypic coefficient of variation (P.C.V) provides an index of phenotypic variability and is estimated by

$$\text{P.C.V} = \frac{\hat{\sigma}_p}{\bar{X}} \times 100 \tag{3.4}$$

Where $\hat{\sigma}_p$ is the estimate of phenotypic standard deviation of the variable X and $\bar{X}$ is the mean of the observations of the variable X

3.1.2. GENOTYPIC COEFFICIENT OF VARIATION

Genotypic coefficient of variation (G.C.V) is useful to assess and compare the range of genetic diversity for a quantitative trait and is estimated by

$$\text{G.C.V} = \frac{\hat{\sigma}_g}{\bar{X}} \times 100 \tag{3.5}$$

Where $\hat{\sigma}_g$ is the estimate of genotypic standard deviation of variable X and $\bar{X}$ is the mean of the observations of variable X.

3.1.3. HERITABILITY

The scope for genetic improvement through selection for a particular trait depends on its heritability (h^2)

and in broad sense it is the ratio of the genotypic variance to the phenotypic variance. Heritability is estimated as follows:

$$h^2 = \hat{\sigma}_g^2 / \hat{\sigma}_P^2 \qquad (3.6)$$

3.1.4. PHENOTYPIC CORRELATION

Phenotypic correlation coefficient (r_p) is a measure of the degree of linear relationship between two quantitative traits based on their phenotypic values.

Estimates of this parameter are determined from the part of analysis of covariance table of the relevant characters. Table 3.2 gives a skeleton of the part of the analysis of covariance of a pair of characters X and Y of an experiment in R.B.D. with 't' varieties in 'r' replications.

Table 3.2. Analysis of covariance of the experiment on turmeric with r replications and t varieties.

Source	Degree of freedom	Sum of products	Mean sum of products (M.S.P)	Expected M.S.P
Replications	r-1	S.P.R.(XY)	M.S.P.R.(XY)	
Varieties	t-1	S.P.T.(XY)	M.S.P.T(XY)	$\overline{\sigma}_{e_{XY}} + r\,\overline{\sigma}_{g_{XY}}$
Error	(r-1)(t-1)	S.P.E.(XY)	M.S.P.E (XY)	$\overline{\sigma}_{e_{XY}}$
Total	(r-1)(t-1)			

Phenotypic covariance ($\sigma_{p_{XY}}$), genotypic covariance ($\sigma_{g_{XY}}$) and environmental covariance ($\sigma_{e_{XY}}$) are estimated from the table 3.2 as follows:

$$\hat{\sigma}_{e_{XY}} = M.S.P.E_{(XY)} \qquad (3.7)$$

$$\hat{\sigma}_{g_{XY}} = \frac{M.S.P.T_{(XY)} - M.S.P.E._{(XY)}}{r} \qquad (3.8)$$

$$\hat{\sigma}_{p_{XY}} = \hat{\sigma}_{g_{XY}} + \hat{\sigma}_{e_{XY}} \qquad (3.9)$$

If $\hat{r}_p$ denotes the estimated phenotypic correlation coefficient then.

$$\hat{r}_p = \hat{\sigma}_{p_{XY}} / (\hat{\sigma}^2_{p_X} . \hat{\sigma}^2_{p_Y})^{1/2} \qquad (3.10)$$

Significance of the phenotypic correlation coefficient is tested by using the student's (t) test at error degrees of freedom.

3.1.5 GENOTYPIC CORRELATION

The coefficient of genotypic correlation between two traits measures the degree of association between their genetic values. Estimates of genotypic correlation rests on the resumblance between relatives in a manner analogus to the estimation of heritabilities discussed earlier.

The estimate of genotypic correlation coefficient ($\hat{r}_g$) is given below.

$$\hat{r}_g = \frac{\hat{\sigma}_{g_{XY}}}{\left(\hat{\sigma}^2_{g_X} \cdot \hat{\sigma}^2_{g_Y}\right)^{1/2}} \qquad (3.11)$$

the test of statistical significance of the estimated genotypic correlation coefficient is done by using the student's 't' test with error degree of freedom as indicated in 3.12.

$$t = \frac{|\hat{r}_g|}{\left(\sigma^2_{r_g}\right)^{1/2}} \qquad (3.12)$$

where

$$\sigma^2_{r_g} = \frac{1}{f+1}\left[\frac{1}{2}\left(1-\hat{r}_g^2\right)^2 - \frac{1}{2}\left(1-\hat{r}_g^{2}\right)\left(\frac{1}{D}-\frac{\hat{r}_p\hat{r}_g}{C}\right) + 4\left(\frac{\hat{r}_g}{D}-\frac{\hat{r}_p}{C}\right)^2 + 2\left(1-\hat{r}_g^2\right)^2\frac{\left(1-\hat{r}_p^{2}\right)}{C^2}\right]$$

[Narain et al.(1979)]

f = error degree of freedom

$\frac{1}{D} = 1/2\left[\frac{1}{h^2_X} + \frac{1}{h^2_Y}\right]$

$C = h^2_X \cdot h^2_Y$; h^2_X, h^2_Y are the heritability estimates of the characters X and Y respectively.

3.1.6 GENETIC ADVANCE

Genetic advance (G.A) is a measure of the improvement in the genotypic value in the new population as contrasted to the base population through selection.

genetic advance for the individual character is estimated by using the formula suggested by Lush (1949).

$$G.A. = kh^2\sigma_P \qquad (3.13)$$

where k is the intensity of selection at the desired probability level.

3.1.7. GENETIC GAIN

Genetic gain (G.G.) through selection for the individual character is estimated by using the formula.

$$G.G. = \frac{G.A}{\bar{X}} \times 100 \qquad (3.14)$$

3.2. SELECTION OF CHARACTERS

3.2.1. PRINCIPAL COMPONENT ANALYSIS

In many of the experimental situations of multivariate data analysis the characters may be intercorrelated. In such situations, in order to minimise the effort and also to make the study more systematic, it is always profitable to concentrate on those linear combinations of the variables, which are mainly responsible for the total divergence.

Principal component analysis (PCA) consists in transforming the original set of variables $\underline{X}$ to a new set of derived variables $\underline{Z}$ by the orthogonal transformation.

$\underline{Z} = \underset{\sim}{A}\,\underline{X}$ where $\underset{\sim}{A}$ is the P x P matrix of weighting coefficients. The new variables (components or factors) can be considered as providing a description of the 'structure' of the original set of variables.

The first principal component is that weighted combination of the several original variables which accounts for a maximum amount of the total variation or individual differences represented in the complete set of original variables. In General, the rth principal component is that weighted combination which of all possible weighted combinations independent of the first (r-1) account for a maximum amount of the remaining variation among the observations of the original data. The properties of statistical orthogonality and maximisation of variance uniquely define principal components. The principal components indicate the intrinsic relationships with in a single set of variables and are helpful in reducing the dimensionality of the set of variables without as little loss of information as possible.

The variance associated with a principal component is the characteristic root λ_i. The components are generated in decreasing order of variance.

Suppose $X_1, X_2, \ldots\ldots\ldots X_p$ are a set of random variables of any multivariate normal distribution with mean vector $\underline{\mu}$ and variance covariance matrix $\underset{\sim}{\Sigma}$. The estimate of $\underset{\sim}{\Sigma}$ will be the usual sample variance-covariance matrix $\underset{\sim}{S}$

If $a_{11}, a_{21} \cdot \cdot \cdot \cdot \cdot \cdot a_{P1}$ are weights given to the variables according to their relative importance, the first principal component Z_1 is defined as .

$$Z_1 = \underline{a_1}' \underline{X} \qquad (3.15)$$

The sample variance of these linear combination Z_1 is given by

$$S_{Z_1} = \sum_{i=1}^{P} \sum_{j=1}^{P} a_{ij} S_{ij} a_{ji} = \underline{a_1}' \underset{\sim}{S} \underline{a_1} \qquad (3.16)$$

The first principal component is the vector of weighting coefficients $\underline{a1}$ which maximises the variation subjected to the restriction that sum of squares of $\underline{a_1}'$ is a constant, Say $\sum a_{ij}^2 = 1$ ie $\underline{a_1}' \underline{a_1} = 1$

introducing Lagrange's multiplier $\lambda 1$, a new function is defined which is given below.

$$\phi = S^2_{Z_1} + \lambda 1 \; (\underset{\sim}{I} - \underline{a_1}' \underline{a_1}) \qquad (3.17)$$

$$= \underline{a_1}' \underset{\sim}{S} \underline{a_1} + \lambda 1 \; (\underset{\sim}{I} - a_1' a_1) \qquad (3.18)$$

Maximising ϕ with respect to $\underline{a_1}$

$$\frac{\partial \phi}{\partial \underline{a}_1} = 2 \underset{\sim}{S} \, \underline{a1} - 2 \lambda i \, \underline{a_1} \qquad (3.19)$$

Further $\frac{\partial \phi}{\partial \underline{a_1}} = \underline{0}$ if and only if

$$|\underset{\sim}{S} - \lambda_1 \underset{\sim}{I}| \; \underline{a_1} = 0$$

Equating 3.19 to null vector

$$\underset{\sim}{S} \; \underline{a_1} = \lambda 1 \; \underline{a_1} \qquad (3.20)$$

$$\underline{a_1}' \; \underset{\sim}{S} \; \underline{a_1} = a_1' \lambda 1 \; a_1 \qquad (3.21)$$

$$\underline{a_1}' \; \underset{\sim}{S} \; \underline{a_1} = s^2_{z_1} \qquad (3.22)$$

Hence λ_1 is the characteristic root of the variance-covariance matrix $\underset{\sim}{S}$ and $\underline{a_1}'$ is its associated characteristic vector. The coefficient vector $\underline{a_1}$ must be chosen in such a way that $\lambda 1$ is the greatest characteristic root. Numerous methods exist for the calculation of principal components from covariance or correlation matrices. The Jacobi method, the tridiagonalization method and the iterative method are some of the feasible alternatives to the mathematic solution especially with large matrices. Of these methods, the iterative method is quite simple and straight forward. If the iterated components explain the desired percentage of the total variance, the process can be stopped.

The second principal component vector and the associated variance are obtained in the same way as that described for the first component. With 'n' characters a maximum of 'n' principal component vectors are possible. Some of these variables will add only very little information to the description of the structure and hence need not be retained. A simple and popular method is to ignore the characteristic vectors having characteristic roots less than some pre-defined small value say ($\lambda_i < 1$) After the screening process the number of 'effective components' will be greatly reduced. Significance tests also exist for determining the homogenity of the characteristic roots and the selection of components may be done after applying such preliminary tests of significance.

The principal components partition the total variance of the original variables into 'p' additive components. From the general properties of characteristic roots it follows that,

$$\sum_j \lambda_j = \sum_j S^2_{zj} = \text{Trace } \underset{\sim}{S} \qquad (3.23)$$

The relative contribution of the j^{th} component in the system (E_j) is measured by

$$E_j = \frac{\lambda_j}{\text{Trace } \underset{\sim}{S}} \qquad (3.24)$$

If the principal components are calculated from the correlation matrix, Trace $\underset{\sim}{S}$ is equal to the total number of original characters 'n'. Then the percentage contribution f j^{th} principal component is

$$E_j = \frac{\lambda_j}{n} \times 100 \qquad (3.24a)$$

ust as the Principal components are defined as linear ombinations of the original variables the original varia- les can also be defined in terms of the Principal components.

$$X_1 = C_{11} Z_1 + C_{21} Z_2 + \cdots + C_{p1} Z_p \qquad (3.25)$$

he correlation between j^{th} component and the i^{th} varia- le is given by

$$f_{ij} = \lambda_j^{1/2} a_{ij}/S_i \qquad (3.26)$$

here $S_i = \sqrt{\text{Var } x_i}$ and

$_{ij}$ is the component loading or factor loading. In case rincipal components are calculated from the correlation atrices $S_i = 1$ and

$$f_{ij} = \lambda_j^{1/2} a_{ij} \qquad (3.27)$$

rincipal component analysis using inter-correlation atrix is the same as Factor analysis through principal xis method. The scaling coefficient is $\frac{1}{\sqrt{\lambda_i}}$ and the ector of factor loading $\sqrt{\lambda_i}\ \underline{a_i}$

The component loadings are measures of the relative ontributions of the different variables towards the particular principal component and can be used as measuring rods for the choice of the major characters for the construction of selection models. If a variable has no significant correlation with a component then that variable is not contributing much to the variance of the component. If the variable is correlated with any other component then elimination of the variable is not feasible. Hence screening of the variables has to be done only after extracting the major components. To provide for meaningful representation of the relationship among the original variables, the variables which do or do not relate substantially to each factor should be clearly distinguishable. From the set of original variables only a few could be identified as the major contributors for each component. In situations when ordinary PCA fail to achieve this objective factor axes have to be redefined so that they represent the original variables in a simpler and more meaningful way. The process is termed as normalised varimax rotation of principal components.

In the present study both genotypic correlation matrix and phenotypic correlation matrix were used for the extraction of major components through principal component analysis and factor loadings extracted.

3.2.2. PATH-COEFFICIENT ANALYSIS

Path-coefficient analysis attempts to measure the direct and indirect influences of the different casual variables upon a response variable (effect) and permits the separation of an empirical relationship into a series of component parts with direct and indirect links between them. The experimenter presupposes a closed system of variables where each independent variable affects the dependent variable through a series of Paths so that each correlation is a combination of these paths.

Yield is a complex polygenic quantitative character, greatly affected by environment. Hence, selection of superior genotypes based on yield as such is not effective. For rational improvement of yield, mere identification of the phenotypically and genotypically correlated variables and the estimates of phenotypic and genotypic correlation coefficients often felt inadequate, because they do not provide direct and indirect influences of each of the component characters towards the effect.

In the case of turmeric rhizome yield can be considered as the effect factor on a closed system of "Cause and effect relationship", the causal variables being the yield contributing and morphological characters.

Let the response variable Y be a linear function of the casual factors.

$$Y = \sum_{j=1}^{n} b_i x_i + R \qquad (3.28)$$

Where R is the residual factors.

Correlation between casual variable x_i and the effect factor (yield) Y is given by

$$r_{iY} = \frac{\text{CoV } (x_i, Y)}{(\text{S.D. of } x_i)\ (\text{S.D.of } Y)} \qquad (3.29)$$

$$= \frac{\text{CoV } (x_i, b_1x_1 + b_2x_2 + \ldots. + b_ix_i \ldots + b_nx_n + R)}{(\text{S.D. of } x_i)\ (\text{S.D. of } Y)}$$

$$= b_1\ r_{ji} \frac{\sigma x_1}{\sigma Y} + b_2\ r_{i2} \frac{\sigma x_2}{\sigma Y} + \ldots + b_i \frac{\sigma x_i}{\sigma Y} + \ldots$$
$$+ b_n\ r_{in} \frac{\sigma x_n}{\sigma Y}$$

$$= P_{1Y}\ r_{ji} + P_{2Y}\ r_{i2} + \ldots. + P_{jY} + \ldots. + P_{nY}\ r_{jn}$$

Where $P_{iY} = b_i \frac{\sigma x_i}{\sigma Y}$, the path coefficient of x_i to Y

$$r_{iY} = P_{iY} + \sum_{\substack{j=1 \\ \neq i}}^{n} P_{jY}\ r_{ij} \qquad (3.30)$$

P_{iY} denotes the direct effect of x_i on Y and $P_{jY}\ r_{ij}$ denotes the indirect effect of x_i through x_j on Y.

For the estimation of Path coefficients the following simultaneous equations will be set up.

$$r_{1Y} = P_{1Y} + r_{12}P_{2Y} + \cdots\cdots\cdots + r_{1n}P_{nY}$$

$$r_{2Y} = r_{21}P_{1Y} + P_{2Y} + \cdots\cdots\cdots + r_{2n}P_{nY}$$

$$r_{nY} = r_{n1}P_{1Y} + r_{n2}P_{2Y} + \cdots\cdots + P_{nY}$$

In matrix form

$\underline{A} = \underset{\sim}{B}\ \underline{C}$ Where $\underline{A}$ is nX1 vector of correlations of independent variables with Y, $\underset{\sim}{B}$ is the $n \times n$ matrix of inter correlation coefficients and $\underline{C}$ is the $n \times 1$ vector of Path coefficients.

Path coefficients could be obtained as

$$\underline{C} = \underset{\sim}{B}^{-1}\underline{A} \tag{3.31}$$

The residual effect (h) is estimated as follows:

From 3.28 $Y = b_1X_1 + b_2X_2 + \cdots\cdots + b_nX_n + R$

$$\sigma_Y^2 = \sum_{i=1}^{n} b_i^2\, \sigma_{X_i}^2 + 2 \sum_{i<j} r_{ij} b_i b_j\, \sigma_{X_i}\, \sigma_{X_j} + \sigma_R^2 \tag{3.32}$$

$$1 = \sum_{i=1}^{n} b_i^2 \frac{\sigma_{X_i}^2}{\sigma_Y^2} + 2 \sum_{i<j} r_{ij}\, b_i b_j \frac{\sigma_{X_i}}{\sigma_Y} \frac{\sigma_{X_j}}{\sigma_Y} + \frac{\sigma_R^2}{\sigma_Y^2} \tag{3.33}$$

$$1 = \sum_{i=1}^{n} P_{iY}^2 + 2 \sum_{i<j} r_{ij}\, P_{iY}\, P_{jY} + \frac{\sigma_R^2}{\sigma_Y^2}$$

$$1 = \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}\, P_{iY}\, P_{jY} + \frac{\sigma_R^2}{\sigma_Y^2}$$

$$h^2 = 1 - \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}\, P_{iY}\, P_{jY} \tag{3.34}$$

Where $h^2 = \dfrac{\sigma_R^2}{\sigma_Y^2}$ is the proportional residual variance.

Residual effect h is defined as

$$h = \left(1 - \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} \, P_{iY} \, P_{jY} \right)^{1/2} \qquad (3.35)$$

Residual effect can also be estimated by considering the correlation of Y with Y itself.

$$r_{YY} = b_1 r_{1Y} \frac{\sigma_{X_1}}{\sigma_{Y}} + b2 \; r_{2Y} \frac{\sigma_{X_2}}{\sigma_{Y}} + \ldots\ldots + b_n r_{nY} \frac{\sigma_{X_n}}{\sigma_{Y}} + h^2 \qquad (3.36)$$

$$1 = \sum_{i=1}^{n} P_{iY} r_{jY} + h^2 \quad \text{Where } h^2 = \frac{\sigma_R^2}{\sigma_Y^{\;2}}$$

$$h = \left[1 - \sum_{i=1}^{n} P_{iY} \, r_{iY} \right]^{1/2} \qquad (3.37)$$

In many of the experimental situations the number of characters may be very large. In such situations, all of them need not be included to study the direct and indirect links on the yield. Normally a subset of variables is selected from the total number of variables based on a preliminary correlation analysis. Only those variables which have significant correlations with yield alone are included to study their direct and indirect effects on yield.

3.3. SELECTION INDEX

3.3.1. PHENOTYPIC INDEX

A phenotypic index predicts the expected value of the trait with the help of auxiliary traits.

Basically, the statistical principle of constructing Phenotypic selection indices is the same as that of fitting of a multiple regression equation for predicting the values of a dependent variable from known values of two or more independent variables. The yield of any agricultural crop can be assumed to be a linear function of the effects of a number of independent characters. For this reason the regression analysis find wide applications in the analysis of experiment.

If X_1, X_2...... X_p are the P explanatory variable and Y is the response variable. Multiple regression equation is based on a functional relationship of the form

$$Y = \alpha + \sum_{i=1}^{p} B_i X_i + E \qquad (3.38)$$

where the error components are assumed to be independently and normally distributed with variance 1. The parameters of the fitted equations are estimated by applying the ordinary principle of least squares. If a, b_1, b_2.....b_p are the estimates of α, B_1. . . . . . . B_p then

$$a = \bar{y} - \sum_{i=1}^{P} biXi \qquad (3.39)$$

$$\underline{b} = \underset{\sim}{S}^{-1}\underline{g} \qquad (3.40)$$

Where $\underset{\sim}{S}$ is the P x P matrix of sum of products of traits and $\underline{g}$ is P x 1 vector of sum of products of Xi to y

Coefficient of determination (R^2) is given by

$$R^2 = \frac{\sum bi\ S_{iy}}{S.S.\ y} \qquad (3.41)$$

Where S_iy is the sample covariance of the i^{th} explanatory variable with the dependent variable Y and S.S.y is the total sum of square of Y. R^2 is a measure of the percentage variation in the dependent variable explained by the independent variables of the fitted equation.

The significance of R^2 is tested using the variance ratio tested given by

$$F = \frac{R^2}{1-R^2} \times \frac{n-p-1}{n-p} \qquad (3.42)$$

The adequacy of a restricted simpler model over a full model with large of number of variables is tested using the 'F' test as indicated below.

$$F = \frac{R^2f - R^2r}{1- R^2f} \times \frac{df}{d_r - d_f} \qquad (3.43)$$

Where R_f^2 is the coefficient of determination of the full model, R_r^2 is the coefficient of determination of the restricted model, d_f is the degrees of freedom of R_f^2 and d_r is the degree of freedom associated with R_r^2. A non-significant 'F' test implies that the restricted model is as efficient as the full model in describing the proposed relationship apart from chance variation.

In the present study, a subset of variables alone were retained to build up different prediction equations basing upon the phenotypic performance of the traits.

As a further step attempts were made to identify the best criterion for the selection of characters for building up of the phenotypic selection models. The relative efficiencies of the following six criteria of selection were evaluated.

1. Selection based on the relative magnitudes of the heritability coefficients.
2. Selection based on the relative magnitudes of the genotypic correlation coefficients with yield.
3. Selection based on the relative magnitudes of the phnotypic correlation coefficients with yield.
4. Selection based on the product of heritability and genotypic correlation coefficient with yield.
5. Selection based on the product of heritability and phenotypic correlation coefficient with yield.
6. Selection based on the relative genetic gain of the characters.

A fixed number of important characters was identified based on the above criterion. The character which produced maximum absolute value was the first to be inducted and its contribution towards the total divergence measured in terms of the coefficient of variation. Other characters were also added in the model one after the other depending upon their rank order till the inclusion of a character made no significant improvement in the predictibility of the model.

The most effective criterion for the choice of characters would be the one for which the phenotypic selection model explained maximum amount of variation with the help of the least number of variables.

3.3.2. GENOTYPIC SELECTION INDEX

It was smith (1936) who first applied the theory of discriminant function to develop indices for plant selection. The method of approach suggested is the use of a discriminant function that will best indicate the genotypic value of a plant. The problem is to construct an index based on the phenotypic values which will be used as a basis for selecting out individuals with preferred genetic worth. Smith's method of arriving at the discriminant function is based on maximisation of the correlation between genetic worth and phenotypic performance of various characters.

Smith has defined the genetic worth as,

$$H = \sum_{i=1}^{n} a_i G_i \quad \text{or} \quad H = \underset{\sim}{G}^{\prime}\, \underline{a} \tag{3.44}$$

Where a_i is the relative economic weight and G_i is the genotypic value of the i^{th} character since G_i's are unknown. It cannot be based as a criterion for selection. So selection is to be based upon another function of the observed values of the various characters. This phenotypic function (I) can be assumed to have the following linear from

$$I = \sum_{i=1}^{n} bi\ X_i \quad \text{or} \quad I = \underset{\sim}{X}^{\prime}\underline{b} \tag{3.45}$$

$b_1, b_2 \ldots\ldots\ldots b_n$ are to be estimated such that the correlation between H and I ie $r_{(HI)}$, becomes maximum. Now the variances of the indices are estimated as

$$\sigma_{I}^{2} = \underline{b}^{1}\underset{\sim}{P}\underline{b} \tag{3.46}$$

Where P is the phenotypic variance-covariance matrix and $\sigma_{H}^{2} = \underline{a}^{1}\underset{\sim}{G}\ \underline{a}$ Where $\underset{\sim}{G}$ is the genotypic variance-covariance matrix. The covariance between H and I is estimated as

$$\sigma_{IH} = \underline{b}^{1}\underset{\sim}{G}\ \underline{a} \tag{3.47}$$

The vector of coefficients $\underline{b}$ is to be constructed in such a way that correlation between H and I should be maximum.

$$r_{(HI)} = \frac{\underline{b}^1 \underset{\sim}{G}\, \underline{a}}{(b^1 \underset{\sim}{P}\, b)^{1/2}\ (\underline{a}^1 \underset{\sim}{G}\, \underline{a})^{1/2}} \qquad (3.48)$$

Maximising $\log r_{(HI)}$ is equivalent to maximising $r_{(HI)}$

$$\log r_{(HI)} = \log \underline{b}^1 \underset{\sim}{G}\, \underline{a} - 1/2 \log \underline{b}^1 \underset{\sim}{P}\, \underline{b} - 1/2 \log \underline{a}^1 \underset{\sim}{G}\, \underline{a} \qquad (3.49)$$

Differentiating $\log r_{(HI)}$ with respect $\underline{b}$ and equating to zero we get.

$$\frac{\partial \log r_{(HI)}}{\partial \underline{b}} = \frac{\underset{\sim}{G}\, \underline{a}}{\underline{b}^1 \underset{\sim}{G}\, \underline{a}} - \frac{1}{2} \frac{\underset{\sim}{P}\underline{b}}{\underline{b}^1 \underset{\sim}{P}\, \underline{b}} \qquad (3.50)$$

equating to zero

$$\underset{\sim}{G}\underline{a}. \frac{\underline{b}\, \underset{\sim}{P}\, \underline{b}}{\underline{b}^1 \underset{\sim}{G}\, \underline{a}} = \underset{\sim}{P}\, \underline{b} \qquad (3.51)$$

The scalar $\frac{\underline{b}^1 \underset{\sim}{P}\, \underline{b}}{\underline{b}^1 \underset{\sim}{G}\, \underline{a}}$ can be dropped without affecting the proportionality of b_i's. Therefore

$$\begin{aligned} \underset{\sim}{P}\, \underline{b} &= \underset{\sim}{G}\, \underline{a} \\ \underline{b} &= \underset{\sim}{P}^{-1} \underset{\sim}{G}\, \underline{a} \end{aligned} \qquad (3.52)$$

When $\underline{b}$ arises from $\underline{b} = \underset{\sim}{P}^{-1} \underset{\sim}{G}\, \underline{a}$ the following equalities exist.

$$\sigma_{HI} = \sigma_{I}^{2} \qquad (3.53)$$

$$r_{(HI)} = \sigma_{HI} / \sigma_{I} . \sigma_{H} = \sigma^{2}_{I} / \sigma_{I} . \sigma_{H} \qquad (3.54)$$

$$= \sigma_{I} / \sigma_{H} \qquad (3.55)$$

The expected genetic advance through index selection $(G.A_I)$ can be calculated by using the formula suggested by Robinson et al. (1951)

$$G.A_{I} = K \frac{\sum\sum a_i \; b_i \; g_{ij}}{\sqrt{\sum\sum b_i \; b_j \; p_{ij}}} \qquad (3.56)$$

where K is the intensity of selection at the desired level.

The expected genetic advance through straight selection $(G.A_s)$ is given by

$$G.A_{s} = K \frac{\sum\sum a_i \; a_j \; g_{ij}}{\sqrt{\sum\sum a_i \; a_j \; p_{ij}}} \qquad (3.57)$$

In case rhizome yield (y) alone was considered for straight selection, the expected genetic advance through straight selection for the character can be calculated by the method suggested by Rao (1952). The relevant formula in the context is

$$G.A_{s} = \frac{K \; \sigma^{2}_{g_Y}}{\left(\sigma^{2}_{P_Y}\right)^{\frac{1}{2}}} \qquad (3.57a)$$

where $\sigma^{2}_{g_Y}$ and $\sigma^{2}_{P_Y}$ are genotypic and phenotypic variances of the character (y)

The percentage relative efficiency (R.E) of index selection is given by

$$R.E. = \frac{G.A_I}{G.A_s} \times 100 \qquad (3.58)$$

After constructing the selection indices the varieties are evaluated on the basis of their index scores

The index score for the j^{th} variety is given by

$$V_j = \sum_{i=1}^{p} X_{ij} \; b_i \qquad (3.59)$$

where X_{ij} is the mean of the observations corresponding to the i^{th} character of the j^{th} variety. The varieties are then ranked according to the index scores and the best R% may be retained for further propagation where R is a pre-assigned small quantity.

In order to find out the best system of weighting for constructing the efficient selection model, when proper economic weight of the characters are unknown, selection indices may be evaluated on different criteria such as

(1) Equal economic weights to all characters.

(2) Weights proportional to the heritability estimates

(3) Weights proportional to genotypic correlations with yield.

(4) Weights proportional to phenotypic correlations with yield.

(5) Weights proportional to product of heritability estimates and genotypic correlations with yield.

(6) Weights proportional to product of heritability estimates and phenotypic correlations with yield.

The best system of weighting is seleccted according to the relative magnitudes of the expected genetic advance in each of the above systems.

3.3.3. RESTRICTED SELECTION INDEX

The use of a selection index may sometimes adversely affect the transmission of certain component traits. This situation usually occurs when the direct effects of the traits are high and negative. In such situations the breeder is interested in that selection index, which in addition to resulting in maximum possible increase in the aggregate genotypic economic values, ensures constant mean performance in some of the desired component traits. This is the objective in constructing restricted selection indices as suggested by Kempthorne and Nordskog (1959).

Suppose a breeder wishes to maximise the aggregate genotypic economic values $H = \sum_{i=1}^{n} a_i G_i$ based on 'n' characters, subject to the constraint that the genotypic values for 'r' ($<n$) characters do not change ie. $G_k = V_k$ a fixed quantity for $K = 1,2, \ldots \ldots r$ In the foregoing formulation a_is as before are known economic weights.

Let the linear function of the phenotypic value be given by $I = \sum_{i}^{n} bi\ Pi$. The parameters are estimated in such way that $r_{(HI)}$ is maximum subject to the condition that I and G_k ($K=1,2, \ldots \ldots r$) are uncorrelated.

ie CoV $(I, G_k) = 0$, for $k = 1, 2, \ldots r$

From 3.48

$$r_{(H, I)} = \underline{b}^1 \underset{\sim}{G}\, \underline{a} \Big/ \sqrt{(\underline{a}^1 \underset{\sim}{G}\, \underline{a})\ (\underline{b}^1 \underset{\sim}{P}\, \underline{b})}$$

if $C_1^{\ 1} = (1, 0, 0 \ldots\ldots 0)$

$C_2^{\ 1} = (0, 1, 0 \ldots\ldots 0)$

$C_k^{\ 1} = (0, 0, 0 \ldots\ldots 1)$

then

$$\text{CoV } (I, G_k) = \underline{b}^1 G\, \underline{C_k} = \underline{C}_k^{1}\, G\, \underline{b}, \quad k = 1, 2, \ldots r$$

Maximising $r\ (_{HI})$ subject to the condition

$$\text{CoV } (I, G_k) = 0, \ k = 1, 2, \ldots\ldots r$$

This is equivalent to the maximisation of the quantity

$$\frac{(\underline{b}^1 \underset{\sim}{G}\, \underline{a})^2}{(\underline{a}^1 \underset{\sim}{G}\, \underline{a})\ (\underline{b}^1 \underset{\sim}{P}\, \underline{b})} - 2 \sum_{k=1}^{r} \lambda k\ (\underline{C}_k^1\, \underset{\sim}{G}\, \underline{b}) \qquad (3.60)$$

with regard to $\underline{b}$ where λ_k's are Lagrange multipliers.

This then will lead to the solution set $\underline{b}$ as given below.

$$\underline{b} = (I - \bar{\underset{\sim}{P}}^1 \underset{\sim}{G}\, \underset{\sim}{C}\ (\underset{\sim}{C}^1 \underset{\sim}{G}\, \bar{\underset{\sim}{P}}^1 \underset{\sim}{G}\, \bar{\underset{\sim}{C}}^1)\ \underset{\sim}{C}^1 \underset{\sim}{G})\ \ \bar{\underset{\sim}{P}}^1 \underset{\sim}{G}\, \underline{a} \qquad (3.61)$$

Where I is the identity matrix of order n and c $(C_1\ C_2 \ldots C_r)$ is a matrix of order n x r.

Restricted selection can be applied only after analysing the behaviour of independent characters under study. When the total correlation and direct effect are of the same sign and are nearly equal magnitude then individual selection is beneficial when the total correlation is positive and direct effect is negative then simultaneous selection without any restriction shall be practised. When total correlation is negative and direct effect is positive then simultaneous selection with restriction on the negatively contributed characters has to be employed.

RESULTS

4. RESULTS

The results obtained by the analysis of the data are presented in the tables and diagram appended in this section.

Observations were made on eighteen characters in nineteen varieties of turmeric. The characters studied and symbols used were as follows:

Character	Symbol
1. Length of secondary finger (cm)	X_1
2. Girth of secondary finger at middle (cm)	X_2
3. Number of nodes per secondary finger	X_3
4. Internodal distance of secondary finger (cm)	X_4
5. Length of mother rhizome (cm)	X_5
6. Number of nodes per mother rhizome	X_6
7. Internodal distance of mother rhizome (cm)	X_7
8. Girth of mother rhizome at Centre (cm)	X_8
9. Number of nodes per primary finger	X_9
10 Length of primary finger (cm)	X_{10}
11 Internodal distance of primary finger (cm)	X_{11}
12 Girth of primary finger at centre (cm)	X_{12}
13 Length of last fully opened leaf (cm)	X_{13}

14. Breadth of leaf at centre (cm)	X_{14}
15. Petiole length (cm)	X_{15}
16. Height of the plant (cm)	X_{16}
17. Number of leaves per tiller	X_{17}
18. Yield of raw rhizome per bed (kg/plot)	Y

4.1. GENETIC PARAMETERS

From the analysis of variance both phenotypic and genotypic coefficients of variation were estimated for all the eighteen characters including the yield of turmeric and are presented in table 1. Phenotypic coefficient of variation (P.C.V.) ranged from 8.36 per cent for breadth of leaf at centre to 58.42 per cent for yield of raw rhizome per bed. Genotypic coefficient of variation (G.C.V.) ranged from 3.58 per cent for number of nodes per mother rhizome to 42.29 per cent for yield of raw rhizome per bed.

In all cases genotypic coefficient of variation was found to be lower than the corresponding phenotypic coefficient of variation. This indicated that there had been profound influence of the environment on the tested genotypes. A high genotypic coefficient of variation is an indication of the fixable amount of variability that is presented among the cultivars which can be used with advantage in future plant improvement programmes.

Table 1. Estimates of phenotypic coefficient of variation (P.C.V.), Genotypic Coefficient of variation (G.C.V.), Heritability in the broad sense (h^2) Genetic advance (G.A.) and Genetic Gain (G.G.) under 5% intensity of selection for the eighteen variables in turmeric.

Sl. No.	Characters	P.C.V.	G.C.V.	h^2	G.A.	G.G (%)
1.	Length of secondary finger	21.54	15.09	0.4908	1.3653	21.78
2.	Girth of secondary finger at centre	11.51	7.92	0.4737	0.7241	11.23
3.	Number of nodes per secondary finger	18.89	14.97	0.6277	1.6514	24.43
4.	Internodal distance of secondary finger	12.75	7.03	0.3039	0.0737	7.98
5.	Length of mother rhizome	13.84	10.13	0.5366	1.6686	15.30
6.	Number of nodes per mother rhizome	11.04	3.58	0.1054	0.3287	2.40
7.	Internodal distance of mother rhizome	12.15	3.99	0.1076	0.0214	2.69
8.	Girth of mother rhizome at centre	13.61	11.50	0.7148	3.0243	20.03
9.	Number of nodes per primary finger	10.23	7.04	0.4741	1.0397	9.99
10	Length of primary finger	15.67	11.84	0.5700	1.9140	13.41
11	Internodal distance of primary finger	11.65	8.68	0.5740	0.1341	13.55
12	Girth of primary finger at centre	13.52	8.95	0.4381	1.1137	12.21
13	Length of last fully opened leaf	10.66	7.96	0.5574	7.6328	12.24
14	Breadth of leaf at centre	8.36	4.91	0.3442	0.9631	5.93
15	Petiole length	19.03	12.88	0.4580	5.4194	17.95
16	Height of the plant	20.61	17.26	0.7011	10.1837	29.77
17	Number of leaves per tiller	14.03	10.50	0.5598	0.9214	16.18
18	Yield of raw rhizome per bed	58.42	42.29	0.5240	9.7576	63.06

The genotypic coefficient variation is not the only measure to ascertain the amount of heritable variation. Heritability estimates, genetic advance and genetic gain are also helpful to ascertain the amount of heritable variation. The heritability estimates (h^2) in broad sense were calculated for all the eighteen characters including yield. Heritability coefficients for the different characters are presented in table 1. It can be seen that the estimates ranged from 0.1054 for number of nodes per mother rhizome to 0.7148 for girth of mother rhizome at centre. Among other characters, height of the plant (0.7011) and number of nodes per secondary finger (0.6277) showed relatively high heritabilities whereas number of nodes per mother rhizome (0.1054) and internodal distance of mother rhizome (0.1076) showed low heritability. Estimates of heritability of most of the yield contributing and morphological characters were moderate, fluctuating up and down around 0.50

Genetic advance and genetic gain were estimated for each individual character at 5 per cent intensity of selection and are given in table 1. The highest genetic advance was observed for height of the plant (10.1837) followed by yield of raw rhizome per bed (9.7576) while the lowest genetic advance was obtained for internodal distance of

Table 2. Phenotypic and genotypic correlation coefficients between yield and other characters of turmeric.

Sl. No.	Character	Phenotypic correlation	Genotypic correlation
1.	Length of secondary finger	0.4905*	0.5513*
2.	Girth of secondary finger at centre	0.2956*	0.2150
3.	Number of modes per secondary finger	0.3251*	0.3066
4.	Internodal distance of secondary finger	0.3534*	0.3938
5.	Length of mother rhizome	0.3058*	0.4302
6.	Number of nodes per mother rhizome	0.0892	0.1724
7.	Internodal distance of mother rhizome	0.2937*	0.9237
8.	Girth of mother rhizome at centre	0.5152*	0.6783*
9.	Number of nodes per primary finger	0.7089*	0.8122*
10.	Length of primary finger	0.6943*	0.7016*
11.	Internodal distance of primary finger	0.3557*	0.3252
12.	Girth of primary finger at centre	0.4115*	0.4242
13.	Length of last fully opened leaf	0.7711*	0.8690*
14.	Breadth of leaft at centre	0.3598*	0.5698
15.	Petiole length	0.6771*	0.6672*
16.	Height of the plant	0.5862*	0.6997*
17.	Number of leaves per tiller	0.6762*	0.8143*

* Significant at 5% level

mother rhizome (0.0214). The genetic gain ranged from 2.4 per cent for number of nodes per secondary finger to 63.06 per cent for yield of raw rhizome per bed. Height of the plant (29.77%), number of nodes per secondary finger (24.43%) length of secondary finger (21.78%) and girth of mother rhizome at centre (20.03%) also resulted in relatively high genetic gain.

The extent of association between yield and each of the other characters and also among the characters was further investigated. Phenotypic and genotypic correlation coefficients between yield and each of the other characters are given in table 2. All the characters except the number of nodes per mother rhizome exhibited significant Phenotypic correlations with yield at 5 per cent level. Phenotypic correlations with rhizome yield ranged from 0.0892 for number of nodes per mother rhizome to 0.7711 for length of last fully opened leaf. At the phenotypic level yield showed relatively high correlations with number of nodes per primary finger (0.7089), length of primary finger (0.6943) petiole length (0.6771) and number of leaves per tiller (0.6762).

Genotypic correlations with rhizome yield ranged from 0.1724 for number of nodes per mother rhizome to

0.9237 for internodal distance of mother rhizome. Length of secondary finger (0.5513), girth of mother rhizome at centre (0.6783), number of nodes per primary finger (0.8122) length of primary finger (0.7016), length of last fully opened leaf (0.8690), petiole length (0.6672), height of the plant (0.6997) and number of leaves per tiller (0.8143) were the characters which showed significant genotypic correlations with yield.

Phenotypic and genotypic correlation matrices of the explanatory variables are presented in tables 3 and 4. Among the different combinations of characters length of secondary finger and number of nodes per secondary finger showed maximum phenotypic correlation (0.8273) where as highest genotypic correlation was recorded between length of mother rhizome and girth of primary finger at centre (0.9802).

4.2. SELECTION OF CHARACTERS

4.2.1. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis was conducted using the seventeen biometric characters excluding rhizome yield. Components were extracted from both the phenotypic and genotypic correlation matrices. The characteristic roots less than unity were ignored in both the cases.

Table 5. Principal component analysis based on phenotypic correlation matrix: Vectors of component loadings (F), eigen values of corresponding eigen vectors and percentage variation accounted by each component.

Sl. No.	Character	Vectors of component loadings			
		F_1	F_2	F_3	F_4
1.	Length of secondary finger	0.4356	0.6394	0.5914	0.0824
2.	Girth of secondary finger at centre	0.6003	-0.3684	0.3034	0.4032
3.	Number of modes per secondary finger	0.1969	-0.9012	-0.4962	-0.0234
4.	Internodal distance of secondary finger	0.5053	0.6497	0.2342	0.4025
5.	Length of mother rhizome	0.6376	0.3295	0.4491	0.0775
6.	Number of nodes per mother rhizome	0.2921	-0.5175	0.4861	0.3751
7.	Internodal distance of mother rhizome	0.3380	0.6674	-0.0677	0.4139
8.	Girth of mother rhizome at centre	0.7796	0.2357	0.2891	-0.0182
9.	Number of nodes per primary finger	0.7071	-0.3827	-0.3615	0.2044
10.	Length of primary finger	0.7284	0.3065	-0.5582	0.2044
11.	Internodal distance of primary finger	0.3947	0.7496	-0.5027	0.2316
12.	Girth of primary finger at Centre	0.7093	-0.1256	0.3479	0.3910
13.	Length of last fully opened leaf	0.8105	-0.0280	0.1712	-0.3190
14.	Breadth of leaf at centre	0.6495	0.1209	0.2676	-0.0701
15.	Petiole length	0.6848	-0.3715	0.1040	-0.4484
16.	Height of the plant	0.7889	-0.0907	0.2994	-0.3554
17.	Number of leaves per iller	0.7124	-0.0137	0.0953	0.2704
	Eigen value	6.4274	2.6371	2.2959	1.4253
	Variation explained (%)	37.81	15.51	13.51	8.38

From the principal component analysis of the phenotypic correlation matrix, the first four components alone were retained which accounted for 75.21 per cent of the total variation. The vectors of component loadings, eigen values of the corresponding eigen vectors and percentage variation explained by each component are presented in table 5.

In the principal component analysis based on the genotypic correlation matrix major part of variation (91.74%) was accounted by the first three components and hence it was felt that the saliant features of the analysis could be demonstrated by these three components alone. The vectors of component loadings, eigen values of corresponding eigen vectors and percentage variation explained by each component are presented in table 6. The first component called 'the size component' is a weighted mean of the original measurements and was responsible for 53.76 per cent of the total variation. The second component called 'Shape' component had accounted for 20.73 per cent of total variation. The third component was responsible for 17.25 per cent of total variation. The first component was dominated by internodal distance of mother rhizome, girth of mother rhizome at centre, breadth

Table 6. Principal component analysis based on genotypic correlation matrix: Vectors of component loadings (F), eigen values of corresponding eigen vectors and percentage variation accounted by each component.

Sl. No.	Character	Vectors of component loadings		
		F_1	F_2	F_3
1.	Length of secondary finger	0.2980	0.1789	0.9263
2.	Girth of secondary finger at centre	0.6285	-0.5648	-0.1404
3.	Number of nodes per secondary finger	0.0254	-0.1241	0.9863
4.	Internodal distance of secondary finger	0.5686	0.7341	0.3111
5.	Length of mother rhizome	0.8338	-0.2778	-0.5136
6.	Number of nodes per mother rhizome	0.5593	-0.7921	-0.0506
7.	Internodal distance of mother rhizome	0.9716	0.1479	-0.3442
8.	Girth of mother rhizome at centre	0.9613	0.0839	-0.1531
9.	Number of nodes per primary finger	0.7328	0.0824	0.5746
10.	Length of primary finger	0.6760	0.7026	0.1343
11.	Internodal distance of primary finger	0.3670	0.8822	-0.2386
12.	Girth of primary finger at centre	0.7905	-0.3619	-0.1358
13.	Length of last fully opened leaf	0.8425	0.4230	0.2398
14.	Breadth of leaft at centre	0.9196	0.0203	-0.0921
15.	Petiole length	0.7826	-0.3302	0.2213
16.	Height of the plant	0.9139	-0.3854	0.1079
17.	Number of leaves per tiller	0.8377	0.1939	0.2036
	Eigen value	9.1392	3.5232	2.9330
	Variation explained (%)	53.76	20.73	17.25

of leaf at centre and height of the plant. Internodal distance of primary finger has the largest weighting for the second component followed by internodal distance of secondary finger. The major contributors towards variation on the third component are number of nodes per secondary finger and length of secondary finger.

It is evident from table 6 that the characters were not easily distinguishable on the basis of the component loadings and the selection of a subset of characters of a reasonable size seemed to be fruitless. The normalised varimax rotation which was then attempted for the same purpose also found to be ineffective.

Selection of characters was attempted based on the relative magnitudes of the Component loadings. Characters for which component loading exceeded 0.9 alone were selected from the first vector of component loadings. The characters which was selected based on the first vector are internodal distance of mother rhizome (0.9716), girth of mother rhizome at centre (0.9613), breadth of leaf at centre (0.9196) and height of the plant (0.9139). Characters which possessed maximum component loadings alone were selected for the second and third vectors. The character selected based on the secondvector of component loadings was internodal distance of primary finger (0.8822) while that selected on the basis of third vector was the number of nodes per secondary finger (0.9863). These six selected characters were further used to construct the genotypic selection indices.

4.2.2. PATH-COEFFICIENT ANALYSIS

In order to assess the absolute contribution of each of the component characters towards yield, path-coefficient analysis was performed and the results are presented in table 7. A preliminary screening of characters was attempted and only those characters which had significant genotypic correlations with yield alone were considered for conducting the path-coefficient analysis. The eight characters selected for path-coefficient analysis were length of secondary finger, girth of mother rhizome at centre, number of nodes per primary finger, length of primary finger, length of last fully opened leaf, petiole length, height of the plant and number of leaves per tiller. In table 7 correlations between yield on one hand and the various characters on the other hand have been partitioned into direct and indirect effects of the various characters.

Table 7. Path-coefficient analysis in turmeric. Direct and indirect effects of characters on yield.

Characters	Length of secondary finger	Girth of mother rhizome at centre.	Number of nodes per primary finger.	Length of primary finger.	Length of last fully opened leaf.	Petiole length	Height of the plant.	Number of leaves per tiller	Genotypic correlation with yield
Length of secondary finger	1.6664	0.0829	-1.0312	-0.0961	-0.3587	-0.3349	0.4420	0.1809	0.5513
Girth of mother rhizome at centre.	0.2538	0.5444	-0.7625	-0.1025	-0.4606	-0.8125	1.5607	0.4575	0.6783
No.of nodes per primary finger	1.4130	0.3414	-1.2162	-0.1166	-0.4177	-0.6979	1.1673	0.3389	0.8122
Length of primary finger	0.9199	0.3205	-0.8141	-0.1742	-0.5020	-0.2017	0.7834	0.3698	0.7016
Length of last fully opened leaf	0.9459	0.3968	-0.8038	-0.1384	-0.6320	-0.7715	1.3342	0.5378	0.8690
Petiole length	0.4684	0.3713	-0.7123	-0.0295	-0.4092	-1.1915	1.7864	0.3836	0.6672
Height of the plant	0.4017	0.4635	-0.7743	-0.0744	-0.4599	-1.1610	1.8333	0.4708	0.6997
No.of leaves per tiller	0.5314	0.4392	-0.7265	-0.1136	-0.5991	-0.8058	1.5214	0.5673	0.8143

Residual effect h = 0.6492

Diagonal entries which are underlined indicate direct effects of the relevant factors.

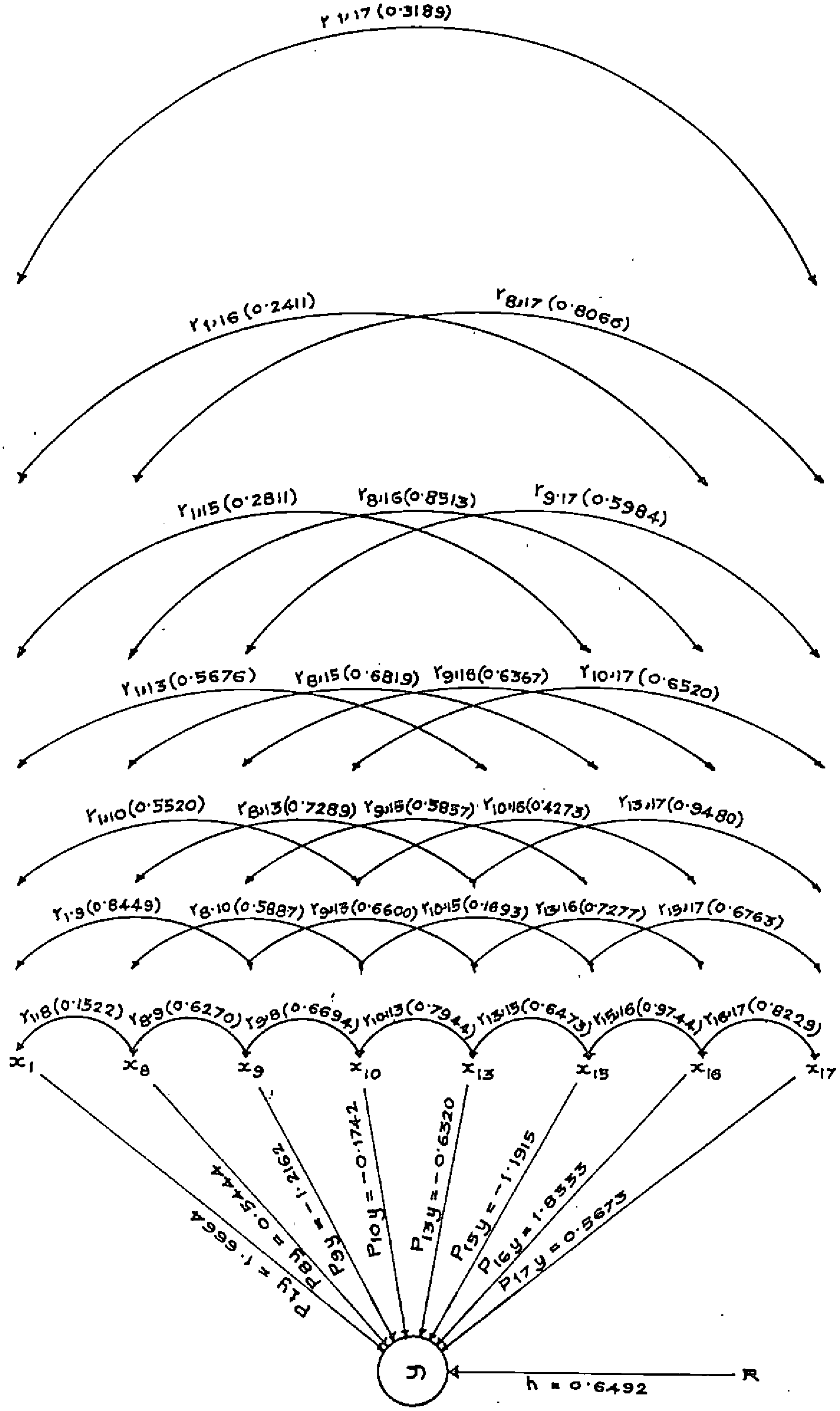

FIG. 1. CAUSE AND EFFECT RELATIONSHIPS BETWEEN YIELD AND OTHER SELECTED BIOMETRIC CHARACTERS.

The cause and effect relationship brought out by the path-coefficient analysis is represented diagramatically in Figure 1.

From table 7 it was evident that though the correlation between length of last fully opened leaf and yield was maximum, it was height of the plant that showed maximum direct effect on yield (1.8333). The other character which exerted high positive direct effect on yield was length of secondary finger (1.6664). The indirect effects of all characters through these traits were positive and high . Direct effects of number of leaves per tiller (0.5673) and girth of mother rhizome at centre (0.5444) on yield were found to be positive. Path coefficient analysis also revealed that number of nodes per finger and petiole length had high negative direct effects (-1.2162 and -1.1915 respectively) on yield. Direct effects of length of last fully opened leaf (-0.6320) and length of primary finger (-0.1742) on yield were found to be negative. The high positive genotypic correlation of these characters with yield were mainly due to their indirect effects through height of the plant and length of secondary finger on yield. The residual effect was estimated to be 0.6492. The relatively high residual effect showed that some of the yield contributing factors had not been taken into account in the path analysis.

Results of path-coefficient analysis indicated that due emphasis would have to be given to height of the plant, length of secondary finger, number of leaves per tiller, girth of mother rhizome at centre, length of last fully opened leaf and length of primary finger for selecting genotypes for further propoagation. Genotypic selection indices were constructed using the above mentioned six characters which included the traits with positive direct effects on yield and those with relatively small negative direct effect but high genotypic correlation with yield.

4.3. SELECTION INDEX

4.3.1. PHENOTYPIC INDEX

Multiple regression analysis was done using the eight characters which exhibited significant genotypic correlation ~~coefficients~~ with yield. A class of multiple linear regression models was fitted to the different sets of character combinations and the coefficients of determination of each model estimated. The tested models and the percentage variation explained by each model are exhibited in table 8. The Coefficient of determination of the various equations ranged from 0.3423 to 0.7888. The superiority of the full model over the simpler models was tested by using the F test.

It was found that there were no significant differences between the parametric models involving seven and six characters respectively. But all the other parametric models

Table 8. Phenotypic indices based on various combinations of characters and the corresponding coefficients of determination (R^2)

Sl. No.	Index label	Characters	Phenotypic index	R^2
1	S_1	X_{16}	$y = -10.1834 + 0.7500\ X_{16}$	0.3423
2	S_2	X_{16}, X_1	$y = -24.4352 + 0.6667\ X_{16} + 2.7282\ X_1$	0.5042
3	S_3	X_{16}, X_1, X_9	$y = -46.4110 + 0.4557\ X_{16} + 0.9960\ X_1 + 3.8494\ X_9$	0.6021
4	S_4	X_{16}, X_1, X_9, X_{15}	$y = -45.7723 + 0.1898\ X_{16} + 0.9794\ X_1 + 3.2635\ X_9 + 0.4856\ X_{15}$	0.6362
5	S_5	$X_{16}, X_1, X_9, X_{15}, X_{13}$	$y = -64.3869 - 0.0888\ X_{16} + 0.1609\ X_1 + 3.0189\ X_9 + 0.3300\ X_{15} + 0.6495\ X_{13}$	0.7304
6	S_6	$X_{16}, X_1, X_9, X_{15}, X_{13}, X_{17}$	$y = -65.5613 - 0.1609\ X_{16} + 0.2398\ X_1 + 2.6936\ X_9 + 0.3345\ X_{15} + 0.5189\ X_{13} + 2.5526\ X_{17}$	0.7565
7	S_7	$X_{16}, X_1, X_9, X_{15}, X_{13}, X_{17}, X_8$	$y = -66.9185 - 0.2070\ X_{16} + 0.2989\ X_1 + 2.6059\ X_9 + 0.3579\ X_{15} + 0.5014\ X_{13} + 2.3624\ X_{17} + 0.3280\ X_8$	0.7595
8	S_8	$X_{16}, X_1, X_9, X_{15}, X_{13}, X_{17}, X_8, X_{10}$	$y = -59.2142 - 0.1822\ X_{16} + 0.1437\ X_1 + 1.1852\ X_9 + 0.5616\ X_{15} + 0.3391\ X_{13} + 1.8734\ X_{17} + 0.1019\ X_8 + 1.6717\ X_{10}$	0.7888

RESULT: Comparison of different phenotypic indices

$S_1\ S_2\ S_3\ S_4\ S_5\ \underline{S_6\ S_7}\ S_8$

were significantly different. The result showed that height of the plant, length of secondary finger, number of nodes per primary finger, petiole length, number of leaves per tiller, girth of mother rhizome at centre, length of last fully opened leaf and length of primary finger are important auxiliary traits which merit phenotypic selection.

To identify the best criterion of selection six groups of phenotypic indices were evaluated in accordance with different criteria of selection viz. (1) heritability (h^2); (2) genotypic correlation with yield (r_g); (3) phenotypic correlation with yield (r_p); (4) product of heritability and genotypic correlation with yield ($h^2 . r_g$); (5) product of heritability and phenotypic correlation with yield ($h^2 . r_p$) and (6) genetic gain (G.G.)

Eight characters were selected on the basis of the highest absolute values for each of the criterion. In any group, the single character index was based on the one having the highest absolute value for the criterion of the choice for that group and characters were added one at a time. Coefficients of determination (R^2) for different combinations of characters selected on the basis of different criteria are given in table 9.

Table 9. Coefficients of determination (R^2) for different combinations of characters selected on the basis of different criteria.

Sl. No.	Criterion of selection					
	h^2		r_g		r_p	
	Characters	R^2	Characters	R^2	Characters	R^2
1	X_8	0.2654	X_7	0.862	X_{13}	0.5923
2	X_8, X_{16}	0.3788	X_7, X_{13}	0.6032	X_{13}, X_9	0.7141
3	X_8, X_{16}, X_3	0.4996	X_7, X_{13}, X_{17}	0.6507	X_{13}, X_9, X_{10}	0.7304
4	X_8, X_{16}, X_3, X_{11}	0.5235	X_7, X_{13}, X_{17}, X_9	0.7491	$X_{13}, X_9, X_{10}, X_{15}$	0.7712
5	$X_8, X_{16}, X_3, X_{11}, X_{10}$	0.6819	$X_7, X_{13}, X_{17}, X_9, X_{10}$	0.7568	$X_{13}, X_9, X_{10}, X_{15}, X_{17}$	0.7881
6	$X_8, X_{16}, X_3, X_{11}, X_{10}, X_{17}$	0.7128	$X_7, X_{13}, X_{17}, X_9, X_{10}, X_{16}$	0.7568	$X_{13}, X_9, X_{10}, X_{15}, X_{17}$ X_{16}	0.7884
7	$X_8, X_{16}, X_3, X_{11}, X_{10}, X_{17}, X_{13}$	0.7421	$X_7, X_{13}, X_{17}, X_9, X_{10}, X_{16}, X_8$	0.7569	$X_{13}, X_9, X_{10}, X_{15}, X_{17}$ X_{16}, X_8	0.7884
8	$X_8, X_{16}, X_3, X_{11}, X_{10}, X_{17}, X_{13}$ X_5	0.7443	$X_7, X_{13}, X_{17}, X_9, X_{10}, X_{16}, X_8, X_{15}$	0.7896	$X_{13}, X_9, X_{10}, X_{15}, X_{17}$ X_{16}, X_8, X_1	0.7886

Contd....

Table 9 (contd.)

Sl. No.	Criterion of selection					
	$h^2 r_g$		$h^2 r_p$		G.G.	
	Characters	R^2	Characters	R^2	Characters	R^2
1	X_{16}	0.3422	X_{13}	0.5923	X_{16}	0.3422
2	X_{16}, X_8	0.3787	X_{13}, X_{16}	0.5989	X_{16}, X_3	0.4412
3	X_{16}, X_8, X_{17}	0.5124	X_{13}, X_{16}, X_{10}	0.6951	X_{16}, X_3, X_1	0.5053
4	$X_{16}, X_8, X_{17}, X_{10}$	0.6652	$X_{13}, X_{16}, X_{10}, X_{17}$	0.7139	X_{16}, X_3, X_1, X_8	0.5358
5	$X_{16}, X_8, X_{17}, X_{10}, X_9$	0.7047	$X_{13}, X_{16}, X_{10}, X_{17}, X_8$	0.7143	$X_{16}, X_3, X_1, X_8, X_{10}$	0.5358
6	$X_{16}, X_8, X_{17}, X_{10}, X_9, X_{15}$	0.7674	$X_{13}, X_{16}, X_{10}, X_{17}, X_8$ X_9	0.7525	$X_{16}, X_3, X_1, X_8, X_{10}, X_{15}$	0.6933
7	$X_{16}, X_8, X_{17}, X_{10}, X_9, X_{15}, X_1$	0.7699	$X_{13}, X_{16}, X_{10}, X_{17}, X_8,$ X_9, X_{15}	0.7884	$X_{16}, X_3, X_1, X_8, X_{10}, X_{15},$ X_{17}	0.6955
8	$X_{16}, X_8, X_{17}, X_{10}, X_9, X_{15}, X_1$ X_5	0.7721	$X_{13}, X_{16}, X_{10}, X_{17}, X_8,$ X_9, X_{15}, X_1	0.7836	$X_{16}, X_3, X_1, X_8, X_{10}, X_{15},$ X_{17}, X_5	0.7007

As expected in all groups coefficient of determination were found to increase with an increase in the number of variables. The model with five characters selected in accordance with the relative magnitudes of phenotypic correlation coefficient explained maximum amount of variation (78.81%). But for the same number of variables percentage variation explained by the other phenotypic indices, with different criteria of selection ranged from 0.6819 to 0.7568. Selection based on genotypic correlation coefficient was the next in efficiency to that based on the phenotypic correlation coefficient and the difference between the two methods was not statistically significant. The genotypic correlation coefficient also can be used for the same purpose. Hence it could be inferred that absolute value of phenotypic correlation with yield is the most effective criterion for choosing characters for constructing phenotypic selection indices.

4.3.2. GENOTYPIC SELECTION INDEX

Using the six characters selected through principal component analysis, genotypic selection indices were constructed for all possible combinations of the characters and expected genetic advance and percentage relative efficiency over straight selection estimated for each of

Table 10. Selection indices for the different combinations of characters selected through Principal component analysis, their respective expected genetic advances and percentage relative efficiencies over straight selection.

Sl. No.	Genotypic selection index	G.A.	Percentage relative efficiency
1	$I = 0.7011X_{16}$	10.1836	105.67
2	$I = 1.4799X_8 + 0.6069X_{16}$	13.6376	141.51
3	$I = 1.7105X_8 - 0.1642X_{14} + 0.6593X_{16}$	14.9197	154.81
4	$I = 0.9548X_3 + 1.6677X_8 - 0.1899X_{14} + 0.6772X_{16}$	15.1010	156.69
5	$I = 0.9507X_3 - 0.0198X_7 + 1.6442X_8 - 0.4339X_{14} + 0.6639X_{16}$	15.1304	157.00
6	$I = 0.8754X_3 - 0.0533X_7 + 1.7431X_8 - 4.3497X_{11} - 0.0026X_{14} + 0.6494X_{16}$	15.1776	157.49

the different indices. From each of the different sets of combinations of characters the most efficient index was identified as the one which resulted in the highest percentage relative efficiency over straight selection. The most efficient indices for different models involving varying number of characters their expected genetic advances and percentage relative efficiencies are given in table 10.

Among the different indices the three parametric index, $I = 1.7105\ X_8 - 0.1642\ X_{14} + 0.6593\ X_{16}$ (4.1) had resulted in relatively high genetic advance (14.9197) and relative efficiency (154.81) with comparatively lesser number of variables. Thus this index can be identified as the best among the class of indices.

The genotypic selection index constructed with all the six characters is indicated in 4.3

$$I = 0.8754X_3 - 0.0533X_7 + 1.7431X_8 - 4.3497X_{11} - 0.0026X_{14} + 0.6494X_{16} \quad (4.2)$$

The expected genetic advance of the index was 15.1776 and the percentage relative efficiency 157.49

In all the above selection indices yield was not taken into account as an explanatory variables. Selection indices were also constructed by incorporating yield also as an additional variable in each of the above selected models. It was found that selection indices constructed by considering yield also as an additional variable were more efficient than those without considering yield. The new selection indices incorporating yield as additional variable, the expected genetic advances of the indices and the percentage of relative efficiencies of index selection over straight selection are given in table 11.

Percentage relative efficiencies of indices ranged from 199.75 to 255.27. Among the selection indices the most efficient index was

$$I = 0.3358\ Y + 2.4911X_8 - 0.4869X_{14} + 0.8422X_{16} \qquad (4.3)$$

The expected genetic advance of the index was 24.1102 and percentage relative efficiency 250.18

A simple index with two variables yield and height of the plant given in 4.4 was also found to be more efficient than direct selection.

$$I = 0.4366Y + 0.9182X_{16} \qquad (4.4)$$

Table 11. Genotypic selection indices for the different combination of characters selected through principal component analysis and yield, their respective expected genetic advances and percentage relative differences over straight selection.

Sl. No.	Genotypic selection index	G.A.	Percentage relative efficiency
1	$I_s = 0.5175y$	9.6369	100.00
2	$I = 0.4366y + 0.9182X_{16}$	19.2505	199.75
3	$I = 0.3448y + 2.1912X_8 + 0.7624X_{16}$	22.8650	237.26
4	$I = 0.3358y + 2.4911X_8 - 0.4869X_{14} + 0.8422X_{16}$	24.1102	250.18
5	$I = 0.2362y + 2.6543X_8 - 0.4986X_{14} + 0.8926X_{16} + 1.8086X_3$	24.5627	254.88
6	$I = 0.2477y - 2.3542X_7 + 2.5613X_8 + 0.0836X_{14} + 0.8509X_{16} + 1.7362X_3$	24.5695	254.95
7	$I_1 = 0.2429y - 3.1423X_7 + 2.7406X_8 + 0.1584X_{14} + 0.9730X_{16} + 7.2360X_{11} + 2.3785X_3$	24.6004	255.27

The expected genetic advance from this index was 19.2505 and percentage relative efficiency 199.75

The genotypic selection index with the selected six characters incorporating yield as an additional variable is indicated in (4.5)

$$I_1 = 0.2429Y - 3.1423X_7 + 2.7406X_8 + 0.1584X_{14} + 0.9730X_{16} + 7.2360X_{11} + 2.3785X_3 \qquad (4.5)$$

The expected genetic advance from I_1 was found to be 24.6004 and percentage relative efficiency 255.27

The characters selected using path-coefficient analysis were also used to construct genotypic selection indices. The characters included in the study were length of secondary finger, girth of mother rhizome at centre, length of primary finger, length of last fully opened leaf, height of the plant and number of leaves per tiller.

Table 12. The most efficient genotypic selection indices for the specific combination of characters selected through path analysis, their respective genetic advances and percentage relative efficiencies over straight selection.

Sl. No.	Genotypic selection index	G.A.	Percentage relative efficiency
1	$I = 0.7011\ X_{16}$	10.1836	105.67
2	$I = 1.4799\ X_{13} + 0.6069\ X_{16}$	16.9245	175.62
3	$I = 2.2322\ X_8 + 0.2894\ X_{13} + 0.7120\ X_{16}$	20.7590	215.41
4	$I = 2.1614\ X_8 + 1.1232\ X_{10} + 0.2096\ X_{13} + 0.7441\ X_{16}$	22.3230	231.64
5	$I = 2.0289\ X_8 + 0.9400\ X_{10} + 0.1098\ X_{13} + 0.6734\ X_{16} + 3.5359 X_{17}$	23.6480	245.39
6	$I = 1.1047\ X_1 + 2.0910\ X_8 + 0.6817\ X_{10} + 0.0960\ X_{13} + 0.6863\ X_{16} + 3.5645\ X_{17}$	24.3200	252.36

Genotypic selection indices were constructed for all combinations of the above characters and genetic advances and percentage relative efficiencies were estimated. The most efficient selection indices corresponding to varying number of character combinations, their expected genetic advances and percentage relative efficiencies over straight selection are given in table 12.

Among the selection indices the most efficient index was

$$I = 1.1047X_1 + 2.0910X_8 + 0.6817X_{10} + 0.0960X_{13} + 0.6863X_{16} + 3.5645X_{17} \qquad (4.6)$$

The expected genetic advance from this index was found to be 24.32 with percentage relative efficiency over straight selection 252.36

A simple index with three variables is as indicated in 4.7

$$I = 2.2322X_8 + 0.2894X_{13} + 0.7120X_{16} \qquad (4.7)$$

For this index expected genetic advance and percentage relative efficiency over straight selection were 20.7590 and 215.41 respectively.

Genotypic selection indices were also constructed by incorporating yield as an additional variable in each

Table 13. The most efficient genotypic selection indices for the specific combination of characters selected through path analysis and yield, their respective genetic advances and percentage relative efficiencies over straight selection.

Sl. No.	Genotypic selection index	G.A.	Percentage relative efficiency
1	$I_s = 0.5185y$	9.6369	100.00
2	$I = 0.4336y + 0.9182X_{16}$	19.2505	199.75
3	$I = 0.4804y + 0.4196X_{13} + 1.0429X_{16}$	26.4418	274.38
4	$I = 0.3877y + 2.8660X_8 + 0.3391X_{13} + 0.8174X_{16}$	30.3444	314.87
5	$I = 0.2382y + 2.6376X_8 + 0.2671X_{13} + 0.7494X_{16} + 4.8088X_{17}$	31.7643	329.61
6	$I = 0.0967y + 2.5903X_8 + 1.4482X_{10} + 0.2823X_{13} + 0.8050X_{10} + 4.8091X_{17}$	33.3554	346.12
7	$I_2 = 0.0306y + 1.8360X_1 + 2.7391X_8 + 1.0325X_{10} + 0.2552X_{13} + 0.8381X_{16} + 4.9643X_{17}$	34.1950	354.84

of the above selection models. It was found that selection indices constructed by considering yield also as an additional variable were more efficient than those without considering yield. The new selection indices incorporating yield as an additional variable, the expected genetic advances and percentage relative efficiencies over straight selection are given in table 13.

The most efficient selection index was

$$I_2 = 0.0360Y + 1.8360X_1 + 2.7391X_8 + 1.0325X_{10} + 0.2552X_{13} + 0.8381X_{16} + 4.9643X_{17} \qquad (4.8)$$

For this index I_2, expected genetic advance and Percentage relative efficiency over straight selection were 24.1950 and 354.84 respectively. A simple and efficient index incorporating yield as an additional variable is

$$I = 0.3877Y + 2.8660X_8 + 0.3391X_{13} + 0.8174X_{16} \qquad (4.9)$$

with genetic advance 30.3444 and Percentage relative efficiency over straight selection 314.87.

On comparing the relative efficiencies of the most efficient selection indices I_1 and I_2, the index I_2 which was constructed using the characters selected

Table 14. Different systems of weighting, corresponding selection indices, their respective expected genetic advances and percentage relative efficiencies

Sl. No.	System of weighting	Selection index	G.A.	Percentage relative efficiency
1	Equal weightage to all characters	$I = 1.1047X_1 + 2.0910X_8 + 0.6817X_{10} + 0.0960X_{13} + 0.6863X_{16} + 3.5045X_{17}$	24.3200	252.4
2	Proportional to heritability	$I = 0.6357X_1 + 1.3873X_8 + 0.3693X_{10} + 0.0188X_{13} + 0.4765X_{16} + 2.1838X_{17}$	15.4694	160.5
3	Proportional to genotypic correlation	$I = 0.7671X_1 + 1.5429X_8 + 0.5353X_{10} + 0.1055X_{13} + 0.4859X_{19} + 2.7810X_{17}$	15.7395	163.3
4	Proportional to phenotypic correlation	$I = 0.6791X_1 + 1.2953X_8 + 0.5046X_{10} + 0.1053X_{13} + 0.4026X_{16} + 2.3897X_{17}$	18.3442	190.4
5	Proportional to product of heritability and genotypic correlation	$I = 0.4461X_1 + 1.0133X_8 + 0.2905X_{10} + 0.0349X_{13} + 0.3363X_{16} + 1.7099X_{17}$	11.1552	119.9
6	Proportional to product of heritability and phenotypic correlation	$I = 0.3899X_1 + 0.8463X_8 + 0.2743X_{10} + 0.0390X_{13} + 0.0279X_{16} + 1.4640X_{17}$	10.0552	104.3

through Path-coefficient analysis was more efficient than the index I_1 which was constructed using the characters selected through Principal component analysis. Thus path-coefficient analysis seemed to be a better method for identifying the major yield components than principal component analysis in constructing selection models.

Selection indices were formed by using various systems of weighting and their relative efficiencies over straight selection were compared. Selection indices based on different system of weighting, their expected genetic advances and percentage relative efficiencies are given in table 14. From the table it could be seen that the system of assigning equal economic importance to all selected characters was found to produce the most efficient indices. The other systems of weighting were found to be inefficient when compared to the system of equal economic weightage.

4.3.3. RESTRICTED SELECTION INDEX

The restricted selection index constructed to the experimental data by restricting the Length of the secondary finger is given below.

The Index fitted was

$$\begin{aligned} I_3 = 0.3134X_1 &+ 2.9415X_8 + 0.2724X_9 - 0.4751X_{10} \\ &- 6.3279X_{13} + 2.9312X_5 + 6.2152X_{16} \\ &+ 3.8216X_{17} \end{aligned} \quad (4.10)$$

The expected genetic advance by restricted selection was also worked out. As expected, there was no gain in genetic advance through restricted selection. The estimates of the genetic advance for various characters expected through restricted selection are as given below.

Character	G.A.
X_1	-0.6282
X_8	0.2130
X_9	2.199
X_{10}	-0.5787
X_{13}	0.2578
X_{15}	0.8235
X_{16}	9.9587
X_{17}	0.3184

Using the selection index I_2 the index score for each variety was determined and the varieties were ranked accordingly. The index scores of different varieties are presented in table 15.

Table 15. Index scores calculated to the different varieties of turmeric

Sl. No. with Rank	Variety	Index score
1	Kuchupudi	155.0629
2	Chayapasupa	154.1830
3	G.L.Puram-II	151.0704
4	Armoor	150.5551
5	Kodur	150.1965
6	Duggirala	148.4279
7	Nandyal	147.5354
8	Amruthapani Kothapeta	145.0680
9	Amruthapani Kothapela CII-317	143.6003
10	Kasturi Tanuka	141.9274
11	Mannuthy local	141.0104
12	Rajapuri	137.2105
13	Vontimitta	131.0362
14	Dindrigam Ca-69	126.1861
15	Duggirala CII-325	125.1062
16	Amalapuram	118.9042
17	G.L.Puram-I	117.8633
18	Tekurpeta	117.6585
19	Armoor C II-324	104.8702

On the basis of the index scores five varieties were identified as promising varieties which constituted the best 25 per cent of the population of varieties. The selected varieties together with their respective index scores are given below.

1.	Kuchupudi	(155.0629)
2.	Chayapasupa	(154.183)
3.	G.L.Puram-II	(151.0704)
4.	Armoor	(150.5551)
5.	Kodur	(150.1965)

DISCUSSION

5. DISCUSSION

It is now well recognised that the low productivity of crops can be greatly improved by selective breeding. Statistical methods play a pivotal role in the evolution of suitable breeding strategies for yield improvement. Informations on the genetic variability, heritability, inter-relationships among the component characters and the extent of genetic divergence among the genotypes are the necessary pre-requisites for any crop improvement programme. The use of selection index in such a breeding plan would serve a two fold purpose namely to bring out simultaneous genetic progress in several characters and to improve yield through selection for relatively more heritable characters. Discriminant functions (selection indices) also serve as rational yard-sticks for the discrimination of desirable genotypes from undesirable ones based on the phenotypic performance of the individuals. When a large number of variables are to be encountered in the study of variability in any plant species and the nature of their inter-relations are unknown, these different variables would make the result ambiguous and confounded. Construction of selection indices also become unnecessarily cumbersome if some of the less important variables were not left out by way of some kind of preliminary screening. Two popular multivariate techniques namely principal component analysis and path-coefficient analysis are generally applied for the identification of important characters that merit selection. The relative discriminating powers of these two techniques with regard to the isolation and identification of important variables have not been assessed so far on the basis of actual field data. Further, attempts to increase rhizome yield in turmeric have been very meagre and for such studies a clear understanding of relevant characters which have profound effect on yield is very essential. The present investigation on the application of statistical techniques to the breeding of turmeric has been undertaken with the objective of evolving a suitable criterion of selection on turmeric and to compare the two widely accepted techniques of multivariate analysis. Viz. Principal component analysis and path-coefficient analysis with respect their discriminating power of identification of pertinent characters.

The experimental data of a varietal trial on turmeric involving 19 varieties conducted at the College of Horticulture, Vellanikkara, Trichur during the year 1977-1978 were utilised for this purpose. The data were subjected to multiple linear regression analysis, principal component analysis and path-coefficient analysis. Genetic parameters were estimated through analyses of variance and covariance techniques and simple correlation analyses. Phenotypic and genotypic selection indices were constructed for different combinations of characters and their relative efficiencies over straight selection evaluated. A comparison of principal component analysis and path-coefficient analysis was attempted with respect to their relative efficiency in diagonising the major yield contributing characters by way of genetic advance to be expected for various combinations of characters. Attempts were also made to suggest suitable criterion for the preliminary screening of variables prior to conduct of any kind of multivariate analysis. The salient results obtained in the investigation are discussed below.

The rhizome yield in turmeric is controlled by a number of morphological and rhizome characters. The identification as well as assessment of the magnitude of association of characters are important in directing the breeding efforts. In the present study efforts were made to elucidate the essential pre-requisites for subsequent investigations in turmeric.

In general the characters exhibited wide range of variability both at phenotypic and genotypic levels. Maximum phenotypic coefficient of variation was displayed by rhizome yield and that was followed by length of secondary finger and height of the plants. The breadth of leaf at centre recorded the minimum phenotypic coefficient of variation. The maximum genotypic coefficient of variation was observed for rhizome yield. It was followed by height of the plant and length of the secondary finger. Genotypic coefficient of variation was negligibly small in the case of number of nodes per mother rhizome, internodal distance of mother rhizome and breadth of leaf at centre. These characters were highly under the influence of the varying environmental conditions. This fact is also evident from their low heritability values. In all cases genotypic coefficient of variation was found to be lesser than the corresponding phenotypic coefficient of variation indicating the profound influence of environment on the phenotypic expression of the quantitative traits.

High heritability estimates were observed for girth of mother rhizome at the centre and height of the plant indicating that these two characters were less influenced by environmental factors and had more amount of fixable genotypic variation. Mohanty (1979) reported a high heritability value for plant height in turmeric. Heritability value was moderately high for number of nodes per secondary finger, length of mother rhizome, length of primary finger, internodal distance of primary finger, length of last fully opened leaf, number of leaves per tiller and rhizome yield while the same was low for number of nodes per mother rhizome and internodal distance of mother rhizome.

High heritability estimates indicate the effectiveness of selection based on good phenotypic performance but does not necessarily mean a high genetic gain for the particular trait. Johnson et al. (1955) have pointed out that high heritability estimates along with high genetic advance were more useful than heritability values alone in predicting the resultant effect of the best individual . High heritability and high genetic gain indicate the presence of additive gene effects (Panse, 1957). Genetic advance has also been estimated for all characters. Among the different characters height of the plant had highest genetic advance followed by rhizome yield, length of last fully opened leaf, petiole length and girth of mother rhizome at centre. Thus it was evident that height of the plant had high heritability compared with high genetic advance and should be used for selection rhizome yield showed high genetic advance with moderate estimate of heritability. Low estimate of genetic advance with low heritability was observed for internodal distance of mother rhizome. Low estimate of genetic advance was also obtained for internodal distance of secondary finger and internodal distance of primary finger suggesting the role played by non-additive variation.

The highest genetic gain was observed for rhizome yield followed by height of the plant. Moderately high genetic gain was observed for number of nodes per primary finger and girth of mother rhizome at centre.

For a better understanding of the inter-relationships of yield and its component characters genotypic as well as phenotypic correlation coefficients have been calculated. In most cases genotypic correlation coefficients were slightly greater than phenotypic correlation coefficients indicating the significant interaction effect of the environment on the phenotype. All the characters except the number of nodes per mother rhizome showed statistically significant positive correlations with yield at the phenotypic level. At the genotypic level also none of the estimates of correlations with yield were found to be negative. But significant correlations were noticed only for eight characters. They include length of secondary finger, girth of mother rhizome at centre, number of nodes per primary finger, length of primary finger, length of last fully opened leaf, petiole length, height of the plant and number of leaves per tiller. Mohanty (1979) reported a positively significant phenotypic correlation between rhizome yield and number of leaves per tiller. Nambiar (1979) also observed a positively significant phenotypic correlation between rhizome yield and plant height. It is also in conformity with findings of Philip (1983). Purewal (1957) and Pillai (1973) also noticed similar findings in colocassia and ginger respectively. Mohanty and Sarma (1979) reported significant positive phenotypic and genotypic correlations between yield and height of the plant in ginger. From preliminary considerations eight characters which showed significant correlations with rhizome yield at genotypic level were selected for the conduct of path-coefficient analysis.

The inter correlations among different characters were also worked out. Some of the characters showed strong positive correlations among themselves leading to the problem of multicollinearity in regression analysis. At the genotypic level the problem of multicollinearity seems to be more acute than that at the phenotypic level.

Most of the characters were found to be highly inter-related among themselves. For instance, plant height, petiole length, length of last fullyopened leaf, internodal distance of mother rhizome, length of mother rhizome, and girth of secondary finger showed strong positive correlations among themselves. These characters tend to be inherited together and a change in one character is expected to bring a corresponding change in the other related characters.

In such situations selection based merely on the relative values of simple correlation coefficients without considering the interaction between the component characters might prove misleading. Hence principal component analysis and path-coefficient analysis were attempted on the same data for the purpose of identifying the pertinent characters which control rhizome yield in turmeric.

Extraction of factors help in identifying variables and grouping them with in a common pattern of variation. Factor analysis through the principal component method was attempted using both phenotypic and genotypic correlation matrices. Factors having eigen value at least unity alone were retained and they were ranked in the order of their eigen values. As suggested by Mahajan *et al.* (1981) all the components which accounted for at least 75 per cent of the total variability were retained and component loadings worked out. From the principal component analysis of the phenotypic correlation matrix the first four components alone were extracted which accounted for 75.21 per cent of the total variation.

In the case of principal component analysis based on genotypic correlation matrix the major part of the variation (91.74%) was accounted by the first three components. Hence the first three components alone can be regarded as having any practical significance with regard to the measurement of diversity at the genotypic level. The component loadings of the different characters with respect to the extracted factors were worked out and used as a basis for identifying significant variables affected by some common factor.

Walton (1972) found a minimum correlation coefficient of 0.45 among traits affected by a common factor.

Asawa (1981) in his studies on factor analysis in chickpea, restricted selection to only those characters which gave a factor loading greater than or equal to 0.5. Ramana Rao *et al.* (1981) selected those characters which gave a minimum correlation of 0.8 with the factor. He also ranked the characters according to their relative loadings and only those characters which ranked first or second on the factor loading scale alone were retained from the second, third and other components of higher order. In the present study most of the characters exhibited very high correlation with extracted components. The relative magnitudes of the component loadings were greater when the analysis was performed in the genotypic correlation matrix than on the phenotypic correlation matrix.

In the case of principal component analysis based on the genotypic correlation matrix, the four characters selected on the first component were those for which component loadings were greater than 0.9 and for the other components the traits which showed maximum component loadings alone were selected. The selected six characters were further used to construct genotypic selection indices.

The first component extracted through principal component analysis based on phenotypic correlation matrix was mainly constituted by morphological and primary finger characters. Mother rhizome characters are of secondary importance to these group of characters in determining the first factor. It was also noticed that the morphological characters were important only in the development of the first component. It indicates their decisive role in controlling the production of the crop. The secondary finger characters were major contributors towards the second component followed by mother rhizome characters and internodal distance of primary finger. The third component was dominated by the length of secondary and primary fingers.

Principal component analysis based on genotypic correlation matrix resulted on the extraction of three major components. The first component was mainly dominated by morphological and mother rhizome characters followed by some of the primary finger characters. The second component was mainly affected by the internodal distance of primary and secondary fingers. It indicates that these two characters are likely to behave alike in inheritance. In a broad sense, the third component could be designated as a primary finger character group. The important character which contributed to variations in the third factor are length of secondary finger and internodal distance of secondary finger. Thus it can be inferred that these two characters are controlled by a common factor and hence they are expected to behave alike in segregation.

The study revealed the importance of morphological characters in controlling crop production. The fact that all the morphological characters share a common pattern of variation along with important yield components rightly exhibit their dominant role in inducing genetic diversity and improving yield. Next to morphological characters some of the rhizome characters also play a vital role in improving yield.

Path-coefficient analysis was also done to determine the direct and indirect effects of the various characters. Length of secondary finger girth of mother rhizome at centre, number of nodes per primary finger, length of primary finger, height of the plant and number of leaves per tiller were found to have significant genotypic correlations with yield. Hence, they were selected for the study of path-coefficients.

In the present study it was observed that maximum positive direct effect was contributed by height of the plant towards the rhizome yield. This result is in full agreement with the findings of Nambiar (1979) but in disagreement with that of George (1981). In ginger, Ratnambal (1979) reported high positive direct effect of plant height on yield.

The high positive direct effect of plant height on rhizome yield seemed to be diminished by the negative indirect effects through petiole length and number of nodes per primary finger. As expected, in this investigation the direct effect of length of secondary finger was found to be very high. This high direct effect of length of secondary finger on rhizome yield was diminished mainly by negative indirect effect through number of nodes per primary finger. The direct effects of number of leaves per tiller and girth of mother rhizome at centre were positive.

The increased height of the aerial shoot may be helpful for better exposure of the leaves to the sun thereby increasing the photosynthetic efficiency of the plants which accounted for higher yield. The length of secondary finger naturally enhanced the rhizome yield. Number of leaves per tiller also increases the photosynthetic area, leading to enhanced rhizome yield. An increased girth of mother rhizome at centre followed by a corresponding increased rhizome weight in mother rhizome also contributes to better yield.

Number of nodes per primary finger, petiole length, length of last fully opened leaf and length of primary finger were exhibited negative direct effects though the corresponding genotypic correlations with rhizome yield were high and positive. The high genotypic correlations are due to the high indirect effects of characters through plant height and length of secondary finger. George (1981) reported that length of primary finger had negative direct effect on rhizome yield in turmeric.

The characters which had positive direct effects on yield and also those having small negative direct effect but high genotypic correlation with yield could be regarded as the major yield contributors. Such a selection procedure of characters has been adopted by Usharani and Rao (1981) in blackgram and Boomikumaran and Rathinam (1981) in green gram.

In the present study six characters namely height of the plant, length of secondary finger, number of leaves per tiller, girth of mother rhizome at centre, length of primary finger and length of last fully opened leaf were selected as major contributors towards the rhizome yield. Selection indices were constructed using the above mentioned six characters to identify superior genotypes for further propagation.

A class of phenotypic selection indices were constructed through multiple regression analysis using the characters which showed significant genotypic correlations with rhizome yield. Such an approach was used by Narain and Mishra (1976) who showed that the genetic improvement expected on the basis of such selection procedure could be substantial over direct selection procedure if the phenotypic and genotypic correlations between main and auxiliary traits are of opposite signs. Unlike the case of genotypic index a knowledge of the estimates of genetic parameters is not necessary to adopt this procedure. In the present study different phenotypic indices were constructed with various combinations of characters and their relative efficiencies compared. It was found that rhizome yield could be satisfactorily predicted by a linear function of eight auxiliary traits namely height of the plant, length of secondary finger, number of nodes per primary finger, petiole length, number of leaves per tiller, girth of mother rhizome at centre, length of last fully opened leaf and length of primary finger with a reasonable degree of accuracy. The linear model with the above mentioned eight characters was significantly different from all other simpler linear models. Exclusion of girth of mother rhizome at centre from the linear model with seven auxiliary traits had not resulted in any substantial change in the predictability of the equation.

In this study all the variables were not included for the conduct of path-coefficient analysis. A preliminary selection of variables was attempted on the basis of the relative strengths of their genotypic relations with yield. In order to examine the validity of such a selection procedure an attempt was made to compare the relative efficiencies of various selection criteria for effective preliminary screening of variables prior to the conduct of any multivariate analytical procedure. It was found that selection based on phenotypic correlation coefficient was resulted in relatively higher value of coefficient of determination with the smallest number of variables. Hence selection of characters based on phenotypic correlation coefficient is found to be the most effective criterion in selecting the important variables. Selection based on genotypic correlation coefficient has come next in efficiency to that based on the phenotypic correlation coefficient and the difference between the two methods was negligibly small. Thus genotypic correlation coefficient shall also be used for the selection of characters.

Genotypic selection indices were constructed using the six characters selected through the principal component analysis of the genotypic correlation matrix.

The expected genetic advance and percentage relative efficiency of index selection over straight selection were estimated for each of the indices. Estimated genetic advance of index selection using height of the plant alone as the explanatory variable was found to be higher than that through straight selection. Selection indices were also constructed by incorporating yield also as an additional variable. The inclusion of yield for which selection is being done with any other single character or characters increased the efficiency of index selection. The result is in agreement with the findings of Johnson et al.(1955) in soyabean, Miller et al. (1958) in upland cotton and Swarup and Changale (1962) in Sorghum.

All the genotypic selection indices based on single characters other than height of the plant were less efficient than that obtained by straight selection. Among the selection indices constructed the most efficient index is given below.

$$I_1 = 0.2429Y - 3.1423X_7 + 2.7406X_8 + 0.1584X_{14} + 0.9730X_{16} + 7.2360X_{11} + 2.3785X_3$$

The expected genetic advance from this index was 24.6004 and the percentage relative efficiency 255.27

The characters selected using path-coefficient analysis were also utilised to construct genotypic selection indices with and without incorporating yield as an additional trait in the linear model. The genetic advance and percentage efficiency of index selection over straight selection were also estimated. The most efficient index was found to be the one which included the traits rhizome yield, length of secondary finger, girth of other rhizome at centre, length of primary finger and number of leaves per tiller. This index is denoted by I_2.

$$I_2 = 0.0360y + 1.8360X_1 + 2.7391X_8 + 1.0325X_{10} + 0.2552X_{13} + 0.8361X_{16} + 4.9643X_{17}$$

The expected genetic advance and percentage relative efficiency over straight selection for this index were 24.1950 and 354.84 respectively. Thus the index I_2 is expected to bring considerable improvement in the subsequent generations.

On comparing the two indices I_1 and I_2 with regard to their relative efficiency over straight selection, it is evident that I_2 is superior to I_1. Thus it appears that the path-coefficient analysis Out weighs principal component analysis in identifying the major yield contributers. But however this fact in no way limit the utility of principal component analysis as an effective tool for identifying the causative sources of common variation. Principal component analysis is concerned with the structural decomposition of original matrix of inter correlations in terms of particular frame of reference and is limited to the synthesis of casual relationship with a single dependent variable. The advantage of principal component analysis lies in dealing with the problem of linear dependence and the difficulty in interpretability arises only with respect to the labeling of latent variables. Path-coefficient analysis is more alert in identifying the true underlying relationship in cases where the assumption of a closed system is valid or can be taken for granted. Harris (1975) had shown that the method of principal component analysis might some times fail to reflect the true relationship between variables and lead to difficult to interpret results. But it helps to identify the variables which behave alike.

One important problem in construction of genotypic selection indices is to have proper knowledge of the weights to be assigned to the different component characters. This infact depends upon the amount by which each unit of variation in a trait actually increases or decreases the net profit. If the yield components are not directly measurable on an economic scale the problem of assigning relative economic weights to characters becomes complicated. In this study discriminant functions were formed utilising the characters selected through path-coefficient analysis with weights assigned to them in accordance with different criteria such as those proportional to heritability estimates, genotypic correlations with yield, phenotypic correlations with yield, products of heritabilities and genotypic correlations with yield, product of heritabilities and phenotypic correlations with yield and equal economic weights to all the characters. It was found that equal economic weightage to all the characters resulted in the most efficient index. This finding is in contrary with the findings of Rao et al. (1979) in rice who noted a slight advantage for weights proportional to phenotypic correlations and phenotypic path-coefficients over equal economic weightage. It was also found that the absolute value of the product of heritability and phenotypic correlation with yield was the most effective criterion for selecting the important characters to construct efficient genotypic selection indices. This result is in contrary with the findings of Saha and Patnaik (1980) in niger who found out the absolute value of the product of heritability and genotypic correlation with yield as the most effective criterion.

Restricted selection index was also developed by imposing restriction of no change on the genotypic value of length of secondary finger and the expected genetic advance for the different characters were estimated. It was found that the restricted selection index developed by restricting the character length of secondary finger was less efficient than that without restriction on the character.

Index scores were calculated for each of the varieties based on the selection index, I_2 and the topmost five varieties were identified. The selected varieties are Kuchupudi, Chayapasupa, G.L. Puram-II, Armoor and Kodur.

SUMMARY

6. SUMMARY

Investigations were made to formulate suitable selection indices for varietal selection in turmeric based on the characters selected through the techniques of principal component analysis and path-coefficient analysis and to compare the relative discriminating powers of these two widely accepted multivariate techniques for the identification of important yield contributing characters. As a pre-requisite for any crop improvement programme, genetic parameters were also estimated. Attempts were also made to identify an optimum system of weighting in the framing of discriminant functions for selection of genotypes, in case the relative economic values of characters are unknown. Data on 18 biometric characters of 19 varieties of turmeric gathered from a field trial conducted by the Department of Plantation Crops, College of Horticulture, Vellanikkara, Trichur during the year 1977-1978 were utilised for the study. The salient results of the investigation are summarised below.

The variability in each character has been substantiated by the estimates of phenotypic and genotypic coefficient of variations. For all characters, genotypic coefficient of variation was lower than phenotypic coefficient of variation. Among the characters, rhizome yield showed maximum phenotypic and genotypic coefficient of variation. High heritability estimates were manifested by girth of mother rhizome at centre and height of the plant. The highest genetic advance with low heritability was obtained for internodal distance of mother rhizome suggesting the role played by non-heritable variation. The highest genetic gain was observed for rhizome yield followed by height of the plant.

From the correlation studies it was found that all the characters except number of nodes per mother rhizome showed significant positive phenotypic correlations with yield. Length of secondary finger, girth of mother rhizome at centre, number of nodes per primary finger, length of mother rhizome, length of last fully opened leaf, petiole length, height of the plant and number of leaves per tiller were the characters which showed significant positive genotypic correlations with yield. The inter correlations among different characters were also worked out. Most of the characters were highly interrelated which led to the problem of multicollinearity.

Principal component analyses were conducted on the phenotypic and genotypic correlation matrices of the seventeen explanatory variables. The first three components of the Principal component analysis based on the genotypic correlation matrix explained a major part of the total variation. Important characters selected from these components are internodal distance of mother rhizome, girth of mother rhizome at centre, breadth of leaf at centre, height of the plant, internodal distance of primary finger and number of nodes per secondary finger. The above mentioned characters were further used to construct the genotypic selection indices.

Path coefficient analysis was also carried out to study direct and indirect effects of each of the selected characters on yield. Those characters which had significant genotypic correlations with yield alone were include for the study. The results of the analysis showed that height of the plant and length of secondary finger exerted high positive direct effects on rhizome yield. Number of leaves per tiller and girth of mother rhizome at centre also showed positive direct effects on rhizome yield. Direct effects of number of nodes per primary finger and petiole length were high but negative. Length of last fully opened leaf and length of primary finger had relatively small negative direct effects on rhizome yield. All the characters had high positive indirect effects through height of the plant and length of secondary finger on rhizome yield. Thus the study emphasised the importance of plant height and length of secondary finger.

Phenotypic indices were constructed using multiple regression technique utilising the eight characters which showed significant genotypic correlations with yield and coefficients of determination for the different models estimated. It was found that rhizome yield could be satisfactorily explained by a linear combination of eight auxiliary traits namely height of the plant, length of secondary finger, number of nodes per primary finger, petiole length, length of last fully opened leaf, number of leaves per tiller, girth of mother rhizome at centre and length of primary finger. To identify the best criterion of selection, six groups of phenotypic indices were evaluated on the basis of the relative values of coefficients of determination for different criteria of selection. From the study it could be inferred that absolute value of phenotypic correlation coefficient with rhizome yield is the most effective criterion for choosing characters for constructing phenotypic indices followed by absolute value of genotypic correlation coefficient with rhizome yield.

Genotypic selection indices were constructed using the characters selected through both principal component analysis and path-coefficient analysis. Genotypic selection indices were also constructed by incorporating yield as an additional variable in all the above indices. The expected genetic advance and percentage relative efficiency of index selection over straight selection were estimated for all the indices. In all cases it was found that index selection was better than straight selection. Further, it was noticed that inclusion of yield as an additional variable to selection indices increased their efficiency considerably.

On compairing the indices developed on the basis of characters selected through principal component analysis and path coefficient analysis with and without incorporating yield into the linear model, the latter index was found to be relatively more efficient than the former index. Hence it appears that path-coefficient analysis is a better method for identifying the major yield contributing characters than the analysis of principal components. Among the class of selection indices, the most efficient index was the one developed on the basis of all the six characters selected through path-coefficient analysis in addition to yield. The expected genetic advance and percentage relative efficiency over straight selection from this index were 34.1950 and 354.84 respectively.

To identify an optimum system of weighting in the framing of discriminant function for plant selection, genotypic selection indices were formed by using various systems of weighting and their percentage relative efficiencies over straight selection were compared. It can be inferred that equal economic weightage to all the selected characters is the best system of weighting to construct genotypic selection indices.

A restricted selection index was also developed by imposing restrictions on the length of secondary finger and it was found that simultaneous selection without restriction was more beneficial than selection with restriction on the length of secondary finger.

Index scores were calculated for all the varieties using the most efficient index and the best five varieties were selected on the basis of the relative magnitudes of the index scores. The promising varieties which could be recommended for further propagation include Kuchupudi, Chayapasupa, G.L.Puram-II, Armoor and Kodur.

REFERENCE

REFERENCES

Anonymous 1983. All India final estimates of chillies, turmeric, arecanut and garlic during 1981-1982. Indian Cocoa arecanut and spices Journal 7: 30.

Agarwal, R., Jain, R.C., Jha, M.P. and Singh, D. 1980. Forecasting of rice yield using climatic variables. Indian J. agric. Sci. 50: 680-684.

Al-jibouri, H.A., Miller, P.A. and Robinson, H.F.1958. Genotypic and environmental variance and covariances in upland cotton cross of interspecific origin. Agron. J. 50: 633-636.

Asawa, B.M. 1981. Factor analysis in chickpea. Indian J. agric. Sci. 51: 156-159.

Asawa, B.M., Asawa, R.K., Srivastav, P.S. and Somavanshi, K.P.S. 1977. Selection indices in sunflower for yield and oil. Indian J. agric. Sci. 47: 505-507.

Badwal, S.S., Gill, K.S. and Singh, H. 1970. Path-coefficient analysis of seed yield in linseed. Indian J. Genet.Pl. Breed. 30: 551-556.

Bartlett, M.S. 1933. On the theory of statistical regression. Proc. Roy. Soc. Edin. 53: 260-283.

Bhagchandani, P.M. and Choudhury, B. 1980. Correlation and Path-coefficient studies in carrot. Indian J. agric. Sci. 50: 663-666.

Boomikumaran, P. and Rathinam, M. 1981. Correlation and Path analysis in greengram. Madras agric. J. 68: 643-647.

Brim, C.A., Johnson, H.W. and Cockerham, C.C. 1959. Multiple selection criteria in soyabean. Agron. J. 51: 42-46.

Burton, G.W. 1952. Quantitative inheritance in grasses. Proc. VI. Int. Grassld. Congr. 1: 277-283.

Butany, W.T., Singh, M. and Mehra, R.B. 1968. Path analysis of yield components in cotton (Gossipium hirsutum). Indian J. Genet. Pl. Breed. 28: 40-48.

Cattel, R.B. 1965. Factor analysis: An introduction to essentials. Biometrics 21: 190-215.

Chandrappa, H.M., Dhulappanavar, C.V. and Goud, J.V. 1978. Discriminant function in triticale. Indian J. agric. Sci. 48: 641-643.

Denis, J.C. and Adams, M.W. 1978. Factor analysis in plant variables related to yield in dry beans. Crop Sci. 18: 74-78.

Dewey, D.R. and Lu, K.H. 1959. A correlation and path-coefficient analysis of components of crested wheat grass seed production. Agron. J. 51: 515-518.

*Fisher, R.A. 1918. The correlation between relatives on the supposition of mendelian inheritance. Trans. Roy. Soc. Edin. 52: 399-433.

*Fisher, R.A. 1954. Statistical methods for research workers. Biological monograph and Manuals. 5: 130-131.

*Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. Ann. Eugen. 7: 179-188.

*Galton, S.F. 1889. Concept of correlation cited by Fisher (1954). pp.209.

George, H. 1981. Variability in the open pollinated progenies of turmeric. M.Sc. Thesis, Kerala Agricultural University, Vellanikkara, Trichur.

Gill, J.L. and Jensen, E.L. 1968. Probability of obtaining negative estimates of heritability. Biometrics 24: 517-526.

Girshic, M.A. 1936. Principal Components. J.Amer. Stat. Assoc.31: 519-528.

Gupta, V.P. 1971. Selection criteria for improving green fodder quality in pearmillet. Indian J. agric. Sci. 31: 283-289.

Harris, D.L. 1964. Expected and predicted progress from index selection involving estimates of population parameters. Biometrics. 20: 46-72.

Harris, R.J. 1975. A primer of Multivariate Statistics. Accademic Press, London. pp.151-154.

Hazel, L.N. 1943. The genetic basis for constructing selection index. Genetics 28: 476-496.

Hazel, L.N. and Lush, J.L.1942. The efficiency of three methods of selection. J. Hered. 33: 393-399.

Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24: 417-441.

James, J.W. 1968. Index Selection with restrictions. Biometrics. 24: 1015-1018.

Johnson, H.W., Robinson, H.F. and Comstock, R.E.1955. Genotypic phenotypic correlations in soyabean and their importance in selection. Agron. J. 47: 477-483.

Johnston, R.J.1978. Multivariate Statistical Analysis in Geography. Longman Inc., New York, pp.127-182.

Jones, W.O. 1959. Cassava research in Nigera. Proc. IV Symp. Int. Soc. Trop. Root Crops.CI.A.R., Cali. Columbia. 139.

*Kaiser, H.F. 1956. The varimax method of factor analysis. Unpublished doctoral dissertation, Univ. California.

Kaiser, H.F. 1958. The varimax criterion for analytic rotation in factor analysis. Psycometrika 23: 187-200.

Kaiser, H.F. 1960. The application of electronic computers to factor analysis. Educ. Psychol. Measurement 20: 141-151.

Kamalam, P., Biradar, R.S., Hrishi, N. and Rajendran, P.G. 1977. Path analysis and correlation studies in sweet potato. J. Root Crops 3: 5-11.

Kempthorne, O. 1957. An introduction to genetic statistics. John wiley and sons. Inc. New York, U.S.A., pp.291-292.

Kempthorne, O. and Nordskog, A.W.1959.Restricted selection indices. Biometrics 15: 10-19.

Khangura, B.S. and Sandhu, R.S.1972. Path analysis in groundnut (Arachis hypogaca L.) Indian J. agric. Sci. 42: 792-795.

Lawley, D.N. 1963. On testing a set of correlation coefficients for equality. Annals. of mathematical statistics 34: 149-151.

*Lawley, D.N. and Maxwell, A.E.1963. Factor analysis as a statistical method. Butterworth. London.

Li, C.C. 1956. The concept of path-coefficient and it's impacts on population genetics. Biometrics 12: 190-210.

Lin, C.Y. and Allaire, F.R.1977. Heritability of a linear combination of traits. Theor. Appl.Genet. 51: 1-3.

Lush, J.L.1949. Animal breeding plans. Lowa state college press. Amas, lowa. U.S.A.

Mahajan, R.K., Rao, A.V. and Dhindsa, K.S. 1981. Principal component analysis some late-maturing varieties. Indian J. agric. Sci. 51: 9-12.

Maity, S. andChatterjee, B.N. 1977. Growth attributes of potato and their interrelationship with yield. Potato Research 20: 337-341.

Meister, E. and Thompson, N.R. 1976. Genotypic environmental and allometric control of yield components in potato progenies. American potato J. 53: 365.

Miller, D.A., Williams, J.C., Robinson, H.F. and Comstock,R.E.1958. Estimate of genotypic and environment variances and covariances in upland cotton and their implication in selection. Agron. J. 50: 126-131.

Mishra, R.S. and Chhonkar, V.S. 1979. Studieson genetic variability in okra (Abelmoschus esculentus (L.) Moench). Indian J. Hort. 36: 434-437.

Mohanty, D.C.1979. Genetic variability and inter-relationship among rhizome yield and yield components in turmeric. The Andhra agric. J. 26: 77-80.

Mohanty, D.C. and Sarma, Y.N. 1979. Genetic variability and correlation for yield and other variables in ginger germplasm. *Indian J. agric. Sci.* **49**: 250-253.

Murthy, B.R. and Arunachalam, V. 1967. Factor analysis of diversity in the genus sorghum. *Indian J. Genet. Pl. Breed.* **27**: 123-136.

Nagabhushanam, G.V.S., Subramanayam, D. and Reddy, N.S. 1982. Character association and path analysis in erect genotypes of groundnut. *The Andhra agric. J.* **29**: 1-5.

*Nambiar, M.C. 1979. Morphological and cytological investigations in the genus Curcuma Linn. Ph.D. Thesis. University of Bombay.

Narain, P. and Mishra, G.K. 1971. The uses of auxiliary traits in increasing the efficiency of selective breeding. *J. Indian Soc. agric. Stat.*

Narain, P., Bhatia, V.K. and Malhotra, P.K. 1979. Hand book of Statistical genetics. Indian agricultural statistics research institute, New Delhi.

Nigam, A.K., Srivastava, V.K., Jain, J.P. and Gopalan, R. 1979. A note on estimation of heritability by regression analysis. *Biom. J.* **21**: 667-673.

Norman, I.J. 1972. Yield components in sugarcane and selected sugarcane population. *Crop Sci.* **12**: 906-908.

Panse, V.G. 1957. Genetic quantitative characters in relation to plant breeding. *Indian J. Genet. Pl. Breed.* **17**: 318-328.

Pearce, S.C. 1967. Multivariate technique of use in biological research. *Expl. Agriculture* **5**: 67-77.

Pesek, J. and Baker, R.J. 1969. Desired improvement in relation to selection indices. *Can J. Plant Sci.* **49**: 803-804.

Philip, J. and Nair, P.C.S. 1983. Morphological and yield characters of turmeric types. *Indian Cocoa Arec. Spi. J. 6*: 61-67.

Pillai, P.K.T. 1973. Germplasm Collection and screening in ginger. *Ann. Prog. Res. CPCRI. Kasargod*, pp.142.

Purewal, S.S. 1957. Effect of various fertilizers in relation to spacing on development of colocassia. *Indian J. Agron. 2*: 245-255.

Pundir, R.P.S. and Rai, B. 1971. Multiple selection criteria in the genetic improvement of Toria populations. *Indian J. Genet. Pl. Breed. 31*: 377-382.

Rao, C.R. 1952. Advanced Statistical methods in biometric research. John Wiley and Sons inc., New York.

Rao, C.R. 1955. Estimation and tests of significance in factor analysis. *Psychometrika 20*: 93-111.

Rao, T.S. 1972. Note on the natural variability for some quantitative and qualitative characters in bhindi. *Indian J. agric. Sci. 42*: 437-438.

Rao, C.V., Narasimhayya, G. and Rao, K.A. 1973. Path analysis in flue-cured tobacco (*Nicotiana tabacun* L.) *Indian J. agric. Sci. 43*: 198-201.

Rao, C.S., Rawla, S. and Rao, M.G.B.K. 1979. Selection based discriminant function with different relative weight in rice crop. *J. Indian Soc. agric. Stat. 31*: 8.

Rao, V.V.R., Jaisvani, B.G. and Asawa, B.M. 1981. Factor analysis in chilli. Indian J. agric. Sci. 51: 225-228.

*Ratnambal, M.J. 1979.Cytogenetical studies in ginger (Zingiber officinale Rosc.) Ph.D. Thesis. University of Bombay.

Robertson, A. 1959. Sampling variance of genetic correlation coefficients. Biometrics 15: 469-485.

Robinson, H.F., Comstock, R.E. and Harvey, P.H. 1951. Genotypic and Phenotypic correlation in corn and their implication in selection. Agron. J. 43: 282-287.

Saha, D.P. and Patnaik, P. 1980. Selection indices in Niger. Indian J. agric. Sci. 50: 914-917.

Sandhu, J.S. and Korla, B.N. 1976. 'Inter relationship between economic characters and selection indices in onion seed crop. Indian J. Hort. 33:171-175.

Shivahare, M.D., Singh, A.B., Chauhan, Y.S. and Singh,P 1975. Path-coefficient analysis of yield components in India mustard. Indian J. agric. Sci. 45: 422-425.

Singh, K.P. and Singh, A.B. 1974. A selection index for yield in Indian mustard(Brassica junea L.(Zern X Cross) Indian J. agric. Sci. 44: 129-131.

Singh, U. and Singh, D. 1974. Application of discriminant function technique for improving the fodder yield in Sorghum. Indian J. agric. Sci. 44: 127-129.

Singh, D., Singh, H.P. and Singh, P. 1976. Pre-harvest forecasting of wheat yield. Indian J. agric. Sci. 46: 445-450.

Singh, A. and Singh, H.N. 1976. Studies on selection indices in chilli. Indian J. agric. Res. 10: 179-181.

* Smith, H.F. 1936. A discriminant function for plant selection. Ann. Eugen. 7: 240-250.

Swarup, V. and Chaugale, D.S. 1962. Studies on genetic variability in Sorghum. II. Correlation of some important quantitative characters contributing towards yield and application of some selection indices for varietal selection. Indian J. Genet. Pl. Breed. 22: 37-44.

Thurstone, L.L. 1931. Multiple factor analysis. Psychol. Rev. 38: 406-427.

Tikka, S.B.S. and Asawa, B.M. 1978. Factor analysis in lentil. Indian J. agric. Sci. 51: 643-646.

Tikka, S.B.S. and Asawa, B.M. 1982. Factor analysis in cowpea. Indian J. agric. Sci. 52: 846-849.

Usharani, Y and Rao, J.S. 1981. Path analysis of yield components in blackgram. Indian J. agric. 51: 378-381.

Verma, G. and Singh, K.P. 1971. Selection based on multiple characters in barnyard millet. Indian J. agric. Sci. 41: 752-757.

Walton, P.D. 1972. Factor analysis of yield in spring wheat. Crop Sci. 12: 731-733.

Warner, J.N. 1952. A method of estimating heritability. Agron. J. 44: 427-430.

Williams, J.S. 1962. The evaluation of a selection index. Biometrics 18: 375-393.

Wright, S. 1921. Correlation and Causation. *J. agric. Res.* *20*: 557-585.

Wright, S. 1968. Evolution and the genetics of population. The University Press, Chicago, London.

* Original not seen.

A COMPARISON OF PATH-COEFFICIENT ANALYSIS AND PRINCIPAL COMPONENT ANALYSIS IN TURMERIC

BY

GEETHA, U.

ABSTRACT OF A THESIS
Submitted in partial fulfilment of the requirements
for the degree
Master of Science (Agricultural Statistics)
Faculty of Agriculture
Kerala Agricultural University

Department of Statistics
COLLEGE OF VETERINARY AND ANIMAL SCIENCES
Mannuthy, Trichur
1985

ABSTRACT

The discriminant function technique was adopted for the construction of suitable selection indices for rhizome yield in turmeric, utilising the characters selected through principal component analysis and path-coefficient analysis and relative discriminating powers of these two multivariate techniques in identifying the pertinent characters was also evaluated. Data on 18 biometric characters of 19 varieties of turmeric gathered from a field trial conducted by the Department of Plantation Crops, College of Horticulture, Vellanikkara, Trichur during the year 1977-1978 were utilised for the study.

Genetic parameters were estimated for all the biometric characters and it was found that height of the plant gave high heritability combined with high genetic advance and genetic gain which revealed the importance of the character in selection. The correlation studies revealed that most of the characters were highly inter correlated among themselves.

When principal component analysis was attempted on the phenotypic correlation matrix, the first four components contributed about 75 per cent towards the

total variation whereas principal component analysis conducted on genotypic correlation matrix the first three components alone explained about 92 per cent of the total variation. The important characters identified from the analysis of principal components using genotypic correlation matrix include internodal distance of mother rhizome, girth of mother rhizome at centre, breadth of leaf at centre, height of the plant, internodal distance of primary finger and number of nodes per secondary finger.

Path-coefficient analysis revealed that height of the plant and length of secondary finger were the major contributers towards ~~the~~ rhizome yield. Direct effects of number of leaves per tiller and girth of mother rhizome were positive where as number of nodes per primary finger and petiole length had high negative direct effects on rhizome yield.

Multiple regression equations fitted for the different combinations of selected biometric characters revealed that height of the plant, length of secondary finger, number of nodes per primary finger, petiole length, number of leaves per tiller, girth of mother rhizome at centre, length of last fully opened leaf and length of primary finger are collectively responsible for a major portion of phenotypic variation in ultimate rhizome yield. It was also found that absolute value of phenotypic correlation with yield was the most effective criterion for choosing characters for constructing phenotypic indices.

Genotypic selection indices were constructed for different combinations of characters through Principal component analysis and path-coefficient analysis separately, with and without including rhizome yield and their relative efficiencies compared. The result indicated that path-coefficient analysis seemed to be superior to principal component analysis in identifying important yield contributing characters. It was also found that the system of assigning equal economic weights to the selected traits was the more efficient than the other systems of weighting.

Index scores for the different varieties were determined from the most efficient genotypic selection index and the best five varieties were identified for further propagation. They are Kuchupudi, Chayapasupa, G.L.Puram-II, Armoor and Kodur.