

Applications of Principal Component Analysis to Horticultural Research

Amy F. Iezzoni

Department of Horticulture, Michigan State University, East Lansing MI 48824

Marvin P. Pritts

Department of Fruit and Vegetable Science, Cornell University, Ithaca, NY 14853

Horticultural researchers often must measure complex traits, such as vigor, reproductive performance, morphology, or adaptability, and develop relationships with treatments or associated variables. Complex traits, however, are a composite of individual traits that often vary together in response to imposed treatments or evolutionary pressures. Identifying a single variable representative of the complex trait may not be possible, so the researcher is faced with the possibility of separately examining many related variables. If the researcher uses univariate statistics to quantify differences or relationships, then the number of separate analyses required will equal the number of individual variables measured. With many types of biological data, however, correlation among variables is common (multicollinearity), and information provided by separate univariate analyses will be redundant.

Multivariate analysis of variance or multiple regression can be used with multicollinear data sets, but one must know the dependence structure of the variables within such data sets before analysis (i. e., dependent and independent variables must be designated to perform the appropriate analysis). In many cases, there is no *a priori* pattern of causality within a data set. It would be useful to choose a technique that identifies patterns in a data set so that new hypotheses of causality could be proposed and tested.

Principal component (PC) analysis is a method that can be used to identify patterns in a data set and to eliminate redundancy in univariate analyses when multicollinear data are involved. PC analysis essentially restructures data sets containing many correlated variables into smaller sets of components of the original variables. Each set is uncorrelated with any other, but components within sets are related. The resulting combinations may suggest a biological meaning for the grouping of variables or their components. In addition, new values can be assigned to each orthogonal set, and these can usually

replace several original correlated variables in univariate statistical procedures. The intent of this paper is to describe PC analysis, not in mathematical terms, since this is found in textbooks (Sneath and Sokal, 1973; Williams, 1976), but in terms that will allow plant scientists to determine the relevance of this method to their own research.

DESCRIPTION OF THE METHOD

Consider a hypothetical data set in which each of eight objects is scored for four variables.

X_1 :	-4	-4	-2	-2	2	2	4	4
X_2 :	-2	-2	-1	-1	1	1	2	2
X_3 :	-5	-3	-3	-1	1	3	3	5
X_4 :	1	1	-1	-1	-1	-1	1	1

If these variables are somehow related, then it is not expedient to use each of them in separate univariate analyses, because the information gained will be partially redundant. However, it is possible to reduce dimensionality by taking advantage of any existing linear relationships within the data. In this example, X_1 is twice the value of X_2 , and both are closely correlated with X_3 ($r = 0.954$). Hence, the information contained in X_1 , X_2 , and X_3 is almost entirely redundant with regard to the variance-covariance structure. A three-dimensional graph reveals that points fall along a single line through space (Fig. 1 bottom).

One could rotate this line in three-dimensional space so that it becomes horizontal. Because X_4 is uncorrelated with any of the other variables, a vertical axis could be used to express X_4 in relation to this new horizontal axis (Fig. 1 top) with little loss of information. Most importantly, the dimensionality of the data set has been reduced from four to two, making patterns of relationship easy to visualize and interpret.

The logic of PC analysis is similar; data are transformed to create new variables and axes that account for more variation in the original data than any single variable alone. With PC analysis, many variables are examined simultaneously and a transformation equation is extracted that represents a major axis containing as much variation (or dimensionality) from the original data set as possible. If several of the variables in a data set

are linearly related (and assuming they are normally distributed), then the cluster of points in hyperspace will resemble an n-dimensional ellipsoidal cloud.

A second transformation equation is then extracted, one that accounts for as much of the remaining variation as possible, with the condition that values from this equation are uncorrelated with values from the first equation, and the axis from this equation is orthogonal with the first. The amount of variation accounted for by the second transformation is always less than the first. If a second set of correlated traits exists in a data set, then the cloud of points in hyperspace would resemble a flattened football (oval, American type), with the first axis passing through the long axis of the football, and the second axis passing through the flattened portion, perpendicular to the first.

Orthogonal axes continue to be extracted until all variation in the data set is accounted for. At the end of a PC analysis, the number

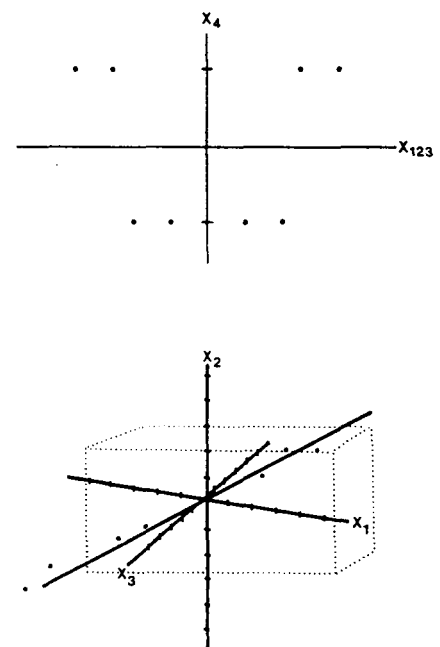


Fig. 1. (Bottom) Vector representing the best-fitting line through three-dimensional space for points contained in the data set. (Top) Values of X_4 plotted against a horizontal axis representing variables $X_{1,2,3}$.

Received for publication 16 Jan. 1990. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked advertisement solely to indicate this fact.

of equations that represent axes will equal the number of original variables, but only the first few equations normally contain useful information.

In the literature, the major axis or vector is usually called principal component 1 (PC1), and subsequent orthogonal axes are numbered sequentially (PC2, PC3, . . . PCn). Each PC axis is defined in terms of a linear transformation of original variable scores. Coefficients of each transformation equation form a set of eigenvectors, and the total variance accounted for by each equation is called the eigenvalue. The sum of the eigenvalues will equal the number of original variables, with each PC accounting for a progressively smaller percentage of the remaining variance.

CONDUCTING THE ANALYSIS

Selecting variables

In a genetic study, a PC analysis is most representative when the cultivars are numerous and unrelated to each other and when many characteristics are evaluated. In other studies, representation from a range of responses or plant types is essential if the interpretation of the analysis is to be meaningful. In either case, variables should be chosen carefully from an array of morphological, structural, reproductive, and growth characteristics.

The characters selected should be of a similar dimension. For example, measurements of volume will have a different variance distribution than measurements of surface area, even though both are measured on the same objects. If measured characters have different dimensionalities, then it is likely that the data are not multivariate normal, so extrapolation of results to a larger population should not be attempted. If the variables will be measured on different scales with greatly different variances, then it is generally advisable to transform the data before analysis by standardizing to zero mean and unit variance. Data sets with both quantitative and qualitative measurements are also not appropriate. Sets of traits that are known to be highly correlated should be avoided since this adds redundant information to the analysis.

Initially, there may be a tendency to mea-

sure many traits; however, after the initial analysis, it may be possible to identify only the most important traits, thereby increasing the efficiency of the method in future applications. Measurements should be made on replicated plots or samples to provide a reliable mean value; however, one must have more observations than variables for the analysis to work. Since it is desirable to have at least twice the number of observations as variables, the technique is most appropriate for survey data.

Input and output

A PC analysis is impractical to conduct with a hand calculator; therefore, computer access is essential. The computer must be instructed to reorganize the data into a variance-covariance matrix or correlation matrix, and use matrix algebra to reduce this data set into a number of linear combinations that form orthogonal axes. With most computer software, one must designate if the inputs are raw variable scores, standardized scores, or a correlation matrix. Options may also exist to assign weights to variables. Advanced applications may even require axis rotation.

The output desired for interpretation and graphing should include 1) the eigenvalue and percentage of total variation explained by each PC, 2) the eigenvectors for each PC (occasionally called "weights"), 3) the PC scores (transformed values for sets of original standardized variables), and 4) the correlations between original standardized variables and the corresponding PC scores (occasionally called "loadings").

INTERPRETING THE ANALYSIS

Interpretation of output from a PC analysis requires that one determine the importance of PCs and variables associated with each PC. There are no rules for making this determination, unlike with univariate statistics, so good judgment must be used.

Two criteria are used to judge whether a PC is meaningful. The first requires an examination of the eigenvalue. If this number is >1.0 , then, theoretically, the corresponding PC has inherently more information than would any single variable alone. All PCs with

an eigenvalue >1.0 would then be subject to interpretation. However, even with random data, at least one PC will have an eigenvalue >1.0 . For this reason, the percentage of total variation explained by the PC is helpful in making a determination. The researcher may have reason to believe that any PC is significant if it explains a certain percentage of total variation in a data set. If biological meaning can be assigned to the PC from an examination of eigenvectors, then confidence in the importance of the PC is increased.

The researcher must use good judgment and knowledge to select the level at which the PC becomes biologically meaningful. The first PC can account for 30% or more of the total variation, even in a completely random data set, depending on the size of the data matrix. Several "Monte Carlo" simulations have been run for the various sized data sets of random numbers to determine the magnitude of their eigenvalues (Karr and Martin, 1981; Stauffer et al., 1985). In addition, Gauch (1982) used artificial data sets to study the relationship among matrix size, inherent variability, and noise reduction. These can be consulted for assistance in judging PC analysis results.

A partial data set by Krahl (1989) can be used to illustrate interpretation of PC analysis. For this data set, six morphological characters were measured on progeny from 11 families of sour cherry (Table 1). A PC analysis indicated that the first PC has an eigenvalue of 4.0 and explains 67.2% of the total variation in the data set (Table 2). This is a relatively large eigenvalue and suggests that PC1 represents the equivalent of four individual variables; therefore, it should be considered for further evaluation. PC2 explains an additional 13.3% of the variation and has an eigenvalue <1.0 . It accounts for a much smaller percentage of the total variation, so it should only be considered significant if some biological meaning can be assigned to it. Additional PCs account for an even smaller percentage of the total variation.

Biological meaning is determined by examining the eigenvectors (weights) and correlations of original variables with PC scores (loadings). Generally, variables with large

Table 1. Full-sib family means of six morphological characters measured on the 'Montmorency', 'Wolynska', and 'English Morello' half-sib families. Abbreviations of the paternal parent in parentheses.

Family	Petiole length (cm)	Leaf length (cm)	Leaf width (cm)	Vein angle (degrees)	Serrations/cm (no.)	Branch angle (degrees)
Montmorency x R1	2.0	10.2	6.4	63	7.9	47
x Meteor Korai	1.3	7.8	4.8	60	8.6	51
x M63 (M6)	1.3	7.8	5.0	55	9.7	50
x Nefris (Nf)	1.5	7.9	4.4	58	10.0	55
x Fruchtbare von Michurin (FM)	1.3	7.8	4.5	60	10.2	47
x R2 (R2)	1.5	7.3	4.5	58	9.8	59
x self (MO)	1.7	7.8	4.4	58	9.0	45
Wolynska x Sumadinka (SU)	1.4	7.7	4.3	57	9.0	60
x Oblacinska (OB)	0.9	5.9	3.6	56	8.4	55
English Morello x Sumadinka (SU)	1.3	8.0	4.7	55	11.1	60
x Oblacinska (OB)	1.1	6.9	3.9	59	8.9	59

weights and loadings, either positive or negative, are considered to be contributors to the PC. In many cases, there is a biological reason why these particular variables are related. In the sour cherry example, petiole length, leaf length, leaf width, and vein angle all have larger weights and loadings than serration density and branch angle. Therefore, the first PC probably represents a latent variable reflecting plant size. Good judgment is again required, because it is not uncommon to find a variable with a large weight and small loading, or vice versa. Situations such as these are common when several variables are highly correlated.

PC2 has a high weight and loading for serration density, so PC2 likely represents this variable alone. The corresponding eigenvalue near 1.0 also suggests that PC2 has the equivalent weight of a single variable. Branch angle is reflected in a similar fashion by PC3. At this point, all the original variables have been strongly associated with a particular PC. PC4 through PC6 have such small eigenvalues that any biological interpretation based on a high weight or loading would have little meaning.

A new "plant size" variable can be derived from the eigenvector of the first PC. To arrive at values (scores) for the new var-

iable, one must use the following transformation equation: $PC1 = W_1(X_1 - \bar{X}_1) + W_2(X_2 - \bar{X}_2) + \dots + W_n(X_n - \bar{X}_n)$, where W_n is the corresponding eigenvector (weight) and X_n is the original variable. If a correlation matrix was used for input, then each product in the transformation equation must be divided by the variance of the corresponding variable.

Insight can often be obtained by plotting the PC scores for individual observations in relation to the important PC axes. Clustering along PC axes suggests that some relationship exists among individuals within a cluster. With sour cherries, the PC scores were plotted with respect to PC1 and PC2 to determine if the resultant pattern revealed anything about the genetic relationships among the 11 families. The family 'Montmorency' × R1 had a high positive value along PC1 because of large leaf and petiole lengths, leaf width, and wide vein angles (Table 2, Fig. 2). The families 'Wolynska' × 'Oblacinska' and 'English Morello' × 'Oblacinska' were most negative along PC1 due to short petiole and leaf lengths and narrow leaf widths. They were separated from their other half-sib families with 'Sumadinka' due to differing values of serration density. The self-pollinated 'Montmorency' family, and the other five

'Montmorency' full-sib families, excluding 'Montmorency' × R1, clustered at the center of the PC. Pollen from R1 was collected in Romania from a clone that has the following pedigree: sour cherry × (sour cherry × sweet cherry). The sweet cherry background obviously contributed the large leaf traits to the progeny, thus positioning it at one extreme of PC1. 'Oblacinska' is a low-vigor, small-leaved sour cherry cultivar grown in Yugoslavia, and these traits were reflected in its progeny and in location of progeny along PC1. In this example, PC analysis clearly was able to separate the germplasm in a way that would be expected, knowing the parental differences.

If clustering does not occur, one cannot conclude that no relationships exist among the variables. PC analysis is sensitive only to linear relationships; higher order relationships may go undetected.

In our example, PC1 accounts for most of the variation in the data set, so in essence, the sour cherry genotypes are separated on the basis of size. If additional variables were measured, such as yield components or physiological traits, [hen PC2 would likely be much stronger and provide further separation on the basis of reproductive behavior or physiological responses that are unrelated to plant size.

PC analysis has many uses other than grouping and clustering. It can be used to quantify variation in different groups of variables. PC scores can also be used in univariate statistics to test for differences in complex traits among groups of individuals, or to measure rates of change through regression analysis. PC analysis can be used to identify variables that contribute little information to underlying relationships; these variables can be deleted from further analyses, resulting in a more simplified data set. PCs can be used in multiple regression analysis to determine the relationship between correlated groups of predictor variables and individual outcome variables. PCs involved in significant relationships can be transformed back into original variables, resulting in a regression equation with fewer predictor variables.

Certain multivariate statistical procedures require an absence of linear dependence among variables so that the covariance matrix can be inverted as part of the analysis. If the matrix cannot be inverted because of multicollinearity, then one could delete in-

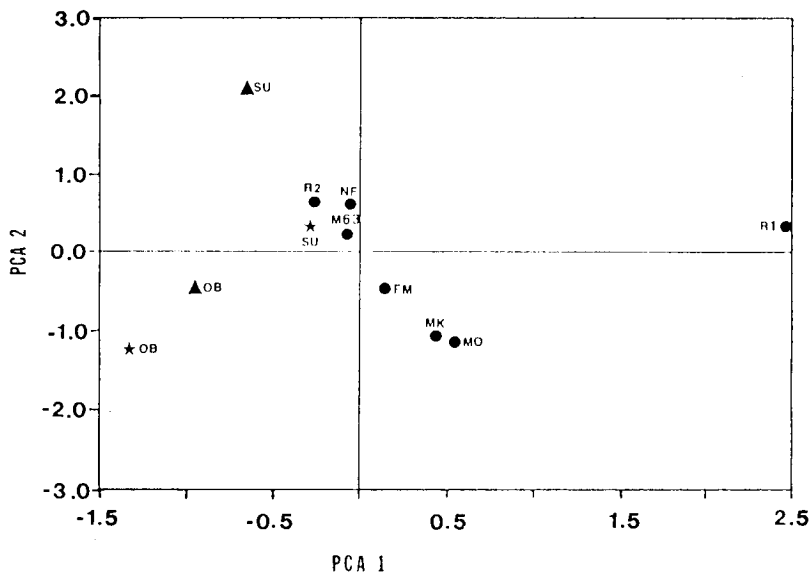


Fig. 2. Positions of PC scores of family means on the first two PC axes for the 'Montmorency', 'Wolynska', and 'English Morello' half-sib families. Maternal parents are indicated by the following symbols: 'Montmorency', ●; 'Wolynska', ★; 'English Morello', ▲.

Table 2. Eigenvectors (W) and loadings (L) of the PC axes from PC analysis of sour cherry families. Eigenvalues and their contribution to total variation are listed at the bottom of columns.

Character	PC 1		PC 2		PC 3		PC 4		PC 5		PC 6	
	W	L	W	L	W	L	W	L	W	L	W	L
Petiole length	0.45	0.88	0.24	0.25	-0.02	0.07	-0.16	-0.05	0.08	-0.38	0.26	0.04
Leaf length	0.46	0.91	0.39	0.34	-0.03	0.11	-0.17	-0.03	-0.17	0.08	-0.76	-0.13
Leaf width	0.46	0.90	0.33	0.27	-0.01	0.08	-0.11	-0.18	-0.57	0.26	0.59	0.09
Vein angle	0.42	0.78	-0.20	-0.37	0.09	0.18	0.89	0.47	0.01	0.04	-0.04	0.02
Serration density	-0.31	-0.40	0.56	0.84	0.74	-0.22	0.20	0.30	0.06	0.02	0.02	0.02
Branch angle	-0.32	-0.64	0.57	0.25	-0.67	-0.72	0.36	0.01	0.05	-0.01	0.03	0.01
Eigenvalue	4.0		0.8		0.65		0.3		0.17		0.03	
Percent of total	67.2		13.3		10.9		5.1		2.9		0.5	

dividual variables until the linear dependency is resolved. An alternative, however, is to perform multivariate procedures on PCs which are, by definition, orthogonal. If PC analysis has been conducted on random samples from a multinormal population whose covariance matrix has a specified structure, then hypothesis tests and confidence intervals for the population can be constructed. These are some of the ways that PC analysis can be used to further our understanding of biological systems. Other examples of PC analysis are illustrated by Broschat (1979). Applications to horticulture are described below.

APPLICATIONS

Yield and quality components

Moore (1965) was one of the first authors to report the use of PC analysis for quantifying yield component interactions in a horticultural crop. Moore wanted to understand the relationship of plant vigor to vegetative growth and to fruit yield in apple. Rather than using a single variable, such as trunk girth, as a measurement of vigor, Moore defined vigor as the sum of numerous growth components, including crop load. He then compared these growth characteristics using PC analysis for the two apple cultivars Cox's Orange Pippin and Worcester Pearmain and identified three major vectors. Vector I, which accounted for 57% and 53% of the variance of 'Cox's Orange Pippin' and 'Worcester Pearmain', respectively, included both the length of shoot growth and the difference in squared girth for the growing season, and was termed the vegetative factor. The second vector, which was independent of the vegetative vigor vector, accounted for 30% of the variance for both cultivars. This vector was identified as the reproductive vector or fruitfulness vector, including such characters as crop weight. The third vector, which accounted for 11% and 13% of the variance for 'Cox's Orange Pippin' and 'Worcester Pearmain', respectively, represented a balance between fruit buds and crop weight. With this interpretation of vigor, it was now possible for the author to apply different treatments to increase yield and to determine whether the yield increase was the result of an increase in vegetative vigor, a change in the balance of vegetative growth and fruit production, a reduction in fruit set, or a combination of these factors.

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

Dale (1988) developed new variables from

PC scores extracted from red raspberry growth and reproductive attributes over several years with genotypes from a half-diallel cross with four parents. General lateral vigor, reproductive vigor, and unachieved reproductive potential were his interpretation of the three major PCs. Dale then performed analysis of variance on these new variables to extract specific and general combining abilities, rather than performing separate statistical analyses on each of the many attributes. Dale also found that general lateral vigor was linearly associated with time of ripening.

Because of the multiplicity of traits used to measure food quality, multivariate statistical procedures, including PC analysis, have been used extensively in food science research (Resurrection, 1988). However, one complex quality trait of importance to horticulturists is an assessment of fruit ripeness or maturity. Fils-Lycaon et al. (1988) used PC analysis to describe the process of fruit ripening in cherry and identify the most important traits that characterized specific stages of ripening. The authors collected data for 21 physiological and chemical characteristics of fruit over a range of harvest dates. PC analysis was able to separate unripe, ripe, and overripe fruit, and the ripe fruit was characterized as having low pH and maximal weight, titratable acidity, and total phenol content.

Taxonomic similarity

The degree of relatedness among individuals within or between species, or among genotypes within a breeding population, can be approximated with PC analysis. The analysis can reveal clusters of morphologically similar phenotypes when plotted against the major PCs. Genotypes within a cluster are likely related. Adams (1977) used PC analysis to calculate the Euclidean distances between cultivars of dry beans along PC axes, and these distances were highly correlated with known genetic relationships based on pedigrees. Such information can then be used to develop hypotheses concerning speciation, to characterize germplasm collections, and in breeding decisions to exploit hybrid vigor or minimize inbreeding depression (Isleib and Wynne, 1983). In addition, covariance of characters may be indicative of genetic linkage or pleiotropy.

This approach has been used infrequently by horticulturists even though it is commonly employed by taxonomists to study species overlap and help explain evolutionary questions (Prance et al., 1969). For example, Jensen and Eshbaugh (1976) examined the relationships between individual oak trees within different populations by means of PC analysis. The authors were able to identify specimens of hybrid origin and their putative parental species.

PC analysis has been used to relate character variation within a species to area of origin (Hillig and Iezzoni, 1988; Hussaini et al., 1977). Recent work by Martin and Adams (1987) illustrates the usefulness of such

an analysis. The authors wanted to investigate the extent and underlying patterns of variability in dry beans when maintained as diverse landraces under the traditional cropping systems of the Malawian farmer. Twenty-five inbred lines grown from seeds collected at each of 15 sites in Malawi were analyzed. Characters evaluated included five phenological and six morphological traits. When plotted on two PCs, five of the landraces from more southerly sites formed a fairly tight cluster, whereas the other nine landraces from the more northerly sites formed a separate, more-dispersed cluster. Seed characteristics varied in a broad north-south cline; however, it was not clear whether environmental factors or traditional consumer preferences were primarily responsible for this gradation. There was no distinct cline for flower color, which may indicate that there has been little or no human or environmental selection for this trait.

Indeterminate (climbing) beans predominated in the northern sites where the farmers frequently grow their beans in association with maize. Phenological characteristics were highly variable within landraces; this may serve to spread out risk during the critical flowering and seed filling periods. The differences among the northerly sites and southernmost sites indicated that some regional adaptation may have occurred among these populations. However, within these regions, considerable diversity exists, suggesting that genetic variability is being maintained as a result of low levels of natural outcrossing between lines of this normally inbreeding species.

Association between genetic and environmental attributes

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

Pritts and Hancock (1984) used PC analysis to consolidate environmental and edaphic variables from sites that supported natural populations of lowbush blueberries. A second PC analysis was also performed to reduce the dimensionality of a group of growth and reproductive attributes that were characteristic of the blueberry populations. The PCs from the environmental variables were then correlated with the PCs from the growth and reproductive variables, resulting in four significant correlations. These correlations among 10 PCs were much easier to interpret than 210 separate correlation coefficients.

Williams et al. (1977) used PC analysis to define three components of growth and yield in sorghum and to correlate these characters with environmental variables. The three

components were defined as crop density, total photosynthate, and vegetative extension. The crop density component was significantly correlated with rainfall in three cultivars, and the total photosynthate component was significantly correlated with maximum temperature. By studying these patterns, the researchers were able to identify germplasm having environmental responses most adapted to their location.

Other

Pritts et al. (1986) tested the hypothesis that foliar applications of 6-benzyladenine (6-BA) to leaves restricts movement of dry matter and mineral elements out of the leaf. Rather than separately analyze the movement of nine mineral elements in strawberry after treatment with various concentrations of 6-BA, they used the first PC as a measure of nutrient status. PC1 accounted for 28% of the variance in nutrient content and eight of the nine nutrients had high weights on the first PC. This PC was linearly related to 6-BA rate, offering support for the hypothesis that 6-BA does restrict movement of mineral elements out of leaves.

SUMMARY

PC analysis reduces a large data set to a small number of unrelated groups of variables or their components, which can be scored along independent, linear axes. Variables strongly associated in the same group may share some underlying biological relationship. These associations are often useful for generating hypotheses or for understanding the behavior of complex traits. The response of groups of variables or complex traits to imposed treatments or evolutionary pressures can be quantified and tested using a combination of PC analysis and simple univariate statistics.

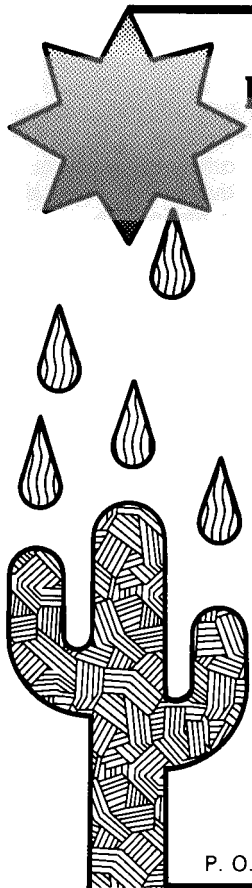
The interpretation assigned to PC groups will depend on 1) the particular set of characteristics to be measured, 2) the diversity of the entries, 3) the strength of underlying relationships among variables, and 4) the judgment of the researcher. Many examples exist in the literature where PC analysis has been used successfully to develop novel hypotheses, to simplify large data sets, or to understand the response of complex traits to imposed treatments or evolutionary pressures.

Literature Cited

- Adams, M.W. 1977. An estimation of homogeneity in crop plants, with special reference to genetic vulnerability in the dry bean, *Phaseolus vulgaris*. *Euphytica* 26:665-679.
- Broschat, T.K. 1979. Principal component analysis in horticultural research. *HortScience* 14:114-117.
- Dale, A. 1988. Relationships between some yield components of the red raspberry (*Rubus idaeus* L.) and the fruiting season: a multivariate approach. *Crop Res.* 28:123-135.
- Fils-Lycaon, B., M. Buret, A. Drouet, C. Hartmann, and F. Duprat. 1988. Ripening and overripening of cherry fruit: Use of principal

- component analysis to check fruit picking and sampling method pertinence and to select the most discriminant analysis criteria. *Sci. des Aliments* 8:383-396.
- Gauch, H.G. 1982. Noise reduction by eigenvector ordinations. *Ecology* 36:1643-1649.
- Hillig, K.W. and A.F. Iezzoni. 1988. Multivariate analysis of a sour cherry germplasm collection. *J. Amer. Soc. Hort. Sci.* 113:928-934.
- Hussaini, S.H., M.M. Goodman and D.H. Timothy. 1977. Multivariate analysis and the geographical distribution of the world collection of finger millet. *Crop Sci.* 17:257-263.
- Isleib, T.G. and J.C. Wynne. 1983. Heterosis in test crosses of 27 exotic peanut cultivars. *Crop Sci.* 23:832-841.
- Jennings, D.L. 1979. The occurrence of multiple fruiting laterals at single nodes of raspberry canes. *New Phytol.* 82:365-374.
- Jensen, R.J. and J.F. Hancock. 1982. Multivariate relationships among California strawberries. *Bul. Torrey Bot. Club* 109:136-147.
- Jensen, R.J. and W.H. Eshbaugh. 1976. Numerical taxonomic studies of hybridization in *Quercus*. I. Populations of restricted areal distribution and low taxonomic diversity. *Systematic Bot.* 1:1-10.
- Karr, J.R. and T.E. Martin. 1981. Random numbers and principal components: Further searches for the unicorn? p. 20-24 In: D.E. Capen (ed.). *The use of multivariate statistics in studies of wildlife habitat*. U.S. Dept. Agr. For. Serv. Tech. Rpt. RM-87.
- Krahl, K.H. 1989. Principal component analysis and isozyme analysis of sweet, sour, and ground cherry germplasm. MS Thesis, Michigan State Univ.
- Martin, G.B. and M.W. Adams. 1987. Landraces of *Phaseolus vulgaris* (Fabaceae) in northern

- Malawi. I. Regional variation. *Econ. Bot.* 41:190-203.
- Moore, C.S. 1965. Inter-relations of growth and cropping in apple trees studies by the method of component analysis. *J. Hort. Sci.* 40:133-149.
- Nevo, E., D. Zahary, A. Brown, and H. Haber. 1979. Genetic diversity and environmental associations of wild barley, *Hordeum spontaneum*, in Israel. *Evolution* 33:815-833.
- Prance, G.T., D.J. Rogers, and F. White. 1969. A taximetric study of an angiosperm family: generic delimitation in the Chrysobalanaceae. *New Phytol.* 68:1203-1234.
- Pritts, M.P. and J.F. Hancock. 1984. Independence of life history parameters in populations of *Vaccinium angustifolium* (Ericaceae). *Bul. Torrey Bot. Club* 111:451-461.
- Pritts, M.P., G.S. Posner, and K.A. Worden. 1986. Effects of 6-BA application on growth and development in 'Tristar', a strong day-neutral strawberry. *HortScience* 21:1421-1423.
- Resurrection, A.V.A. 1988. Applications of multivariate methods in food quality evaluation. *Food Technol.* 40:128-136.
- Sneath, P.H.A. and R.R. Sokal. 1973. *Numerical taxonomy*. Freeman, San Francisco.
- Stauffer, D.F., E.O. Garton, and R. Kirk Steinhorst. 1985. A comparison of principal components from real and random data. *Ecology* 66:1693-1698.
- Williams, W.T. 1976. *Pattern analysis in agricultural science*. Elsevier Scientific, New York.
- Williams, W.T., C.A.P. Boundy, and A.J. Millington. 1977. The effect of sowing date on the growth and yield of three sorghum cultivars in the Ord River Valley. II. The components of growth and yield. *Austral. J. Agr. Res.* 28:381-387.

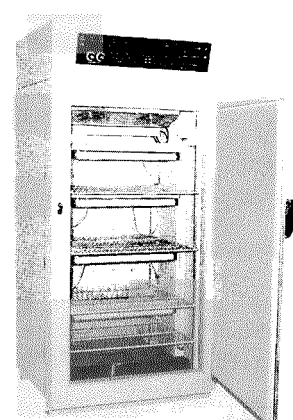



Control the Environment, Effectively and Efficiently

With PERCIVAL

Biological and insect rearing incubators, plant growth and dew chambers. Table models or giant walk-ins. Percival's extensive knowledge and experience is evident throughout our quality line. And Percival flexibility offers the options and modifications for your specific needs.

Write today for our catalog, or call collect 515-432-6501 to discuss your requirements.





Percival Manufacturing Company

P. O. Box 249 • 1805 East Fourth Street • Boone, Iowa 50036