July 1989 A.E. Res. 89-11

A RISK EVALUATION OF GROUNDNUT GENOTYPES IN DROUGHT PRONE AREAS OF INDIA

by
Elizabeth Bailey
and
Richard N. Boisvert

Department of Agricultural Economics
Cornell University Agricultural Experiment Station
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York, 14853

In cooperation with the
International Crop Research Institute
for the Semi-Arid Tropics (ICRISAT)
Patancheru
Andhra Pradesh 502 324
India

A RISK EVALUATION OF GROUNDNUT GENOTYPES

IN DROUGHT PRONE AREAS OF INDIA

by

Elizabeth Bailey and Richard N. Boisvert*

ABSTRACT

A major source of risk facing farmers in the semi-arid tropics is the variation in crop production. Much of this variation is due to fluctuations in environmental conditions affecting yield. One approach to reducing this risk is through the introduction of more stable varieties. The traditional approach to identifying such varieties has important shortcomings.

This study's objectives are to explore ways of generating data on genotype performance, to examine alternative approaches to evaluating genotype performance under variable environmental conditions, and to apply them in selecting risk reducing groundnut genotypes in India.

The performance of rainfed crops is largely determined by the availability of moisture. Historical data on meteorological factors are available for three sites in India. An independent measure of relative water availability is developed. Using data from a single site trial on the response of 22 groundnut genotypes to a range of drought conditions, the yield response to relative water available is modelled. While a number of alternative response functions are tried, the translog is found to fit the response relationship best. The estimated response functions are used to simulate yields for each genotype for a series of years in each location.

Having generated yield distributions, the traditional approach to stability analysis is compared with a number of approaches incorporating economic concepts of risk: mean-variance analysis, ordinary and generalized stochastic dominance, mean-Gini and mean-extended Gini criteria, and the exponential utility, empirical moment generating function approach.

Within each location, results are quite consistent across the different approaches. The degree of risk aversity affects the ordering of genotypes. While high yielding genotypes are preferred over the moderately risk averse range, they are not dominant at higher levels of risk aversion. The incorporation of site specific meteorological information leads to different recommendations than those based on the trial results alone.

Results depend crucially on the estimated response relationships. Given the consistency of results within each location, it seems that the direction of most concern for future research lies less in the choice of selection criteria, than with improved modelling of agro-meteorological relationships.

---

* The authors are Senior Research Officer, Northern Ireland Economic Research Center and Professor, Department of Agricultural Economics, Cornell University, respectively. Support for the theoretical work came from Hatch NY(C)121412. This research would not have been possible without the cooperation of scientists from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Tom Walker alerted us to the need for research of this kind and provided helpful comments on an earlier draft. J. H. Williams and R.C. Nageswara Rao provided the experimental data set.

Table of Contents

List of Tables

List of Figures

Page

A RISK EVALUATION OF GROUNDNUT GENOTYPES IN DROUGHT PRONE AREAS OF INDIA

by

Elizabeth Bailey and Richard N. Boisvert

I - INTRODUCTION

Agricultural producers throughout the world face a variety of price, yield and resource risks which result in year-to-year variability in incomes. The types and severity of the risks depend on the farming system, weather and market conditions and the policy and institutional setting. These risks can be particularly burdensome in less-developed countries. In the semi-arid tropics, a major source of risk facing farmers is the year-to-year variation in crop production. Such variability can be largely attributed to fluctuations in environmental factors that affect plant growth and yield, particularly available moisture. To help stabilize agricultural yields, the international agricultural research centers are working to develop higher yielding varieties that also perform well in variable environments (Hazell 1986). ICRISAT, for example, has an extensive ongoing program to evaluate the performance of groundnut (peanut) genotypes under a range of drought conditions.

The evaluation of varieties has, in general, been based on data from multisite, multiseason nursery trials. The analysis of such data generally follows the approach developed by Finlay and Wilkinson (1963) in which the yields of each genotype at each site are regressed on an environmental index, the mean yield of all genotypes at each site. The slope coefficient is regarded as a measure of a genotype's yield stability. The use of multilocational trial data presumes that the spatial replication reflects the actual distribution of environmental conditions in producing areas. If this presumption is invalid, the evaluations may be of less value to farmers concerned with how a genotype will perform over time in their region. Furthermore, data from such trials are often incomplete and thus, results may be biased.

To deal with the data problems, plant scientists at ICRISAT have undertaken extensive experiments involving 22 groundnut genotypes of comparable maturity, and 96 drought treatments. Results from their statistical and physiological analysis of the data are presented in Nageswara Rao and Williams (1985), Nageswara Rao, Williams and Singh (1985), and Williams *et al.* (1986). The purpose of the research on which this bulletin is based is to extend the analysis and contribute further to the evaluation of the performance of groundnuts in drought prone areas by incorporating economic concepts of risk efficiency and decision making under uncertainty. Yield response functions are estimated from the experimental data. These response functions are unique in that they account for the effects of both the quantity and timing of water application. These are in turn combined with historical meteorological data from three sites in two of India's major groundnut producing regions to simulate yield distributions for each genotype in each location. Having generated yield distributions, the efficient genotypes identified by various risk criteria are compared. The results have implications both for specific genotype selection by location and for the design of future experiments.

The bulletin is organized in the following way. Section *II* contains a brief discussion of the joint regression approach to the analysis of stability used by plant scientists, while section *III* contains a review of the economic concepts of risk and utility maximization, and introduces several alternative approaches that incorporate efficiency criteria and expected utility maximization.

The background to the case study of groundnut improvement in India is given in section *IV*. The three regions included in the analysis and the importance of groundnuts to the Indian economy are described; the details of the experiment conducted by groundnut physiologists at ICRISAT are presented.

Section *V* contains a discussion of the derivation of an independent measure of water stress, relative water availability, and its distribution in the three sample locations. In section *VI*, the response relationships between groundnut yields and relative available water are estimated, and empirical yield distributions are simulated.

The results of alternative approaches to risk analysis are presented in section *VII* and compared with results from the traditional stability analysis of the experimental data. The effects of different assumptions regarding farmers' utility functions and preferences are discussed. Section *VIII* contains conclusions and recommendations for future plant breeding methodology and risk analysis.

II - GENOTYPE-ENVIRONMENT INTERACTION AND STABILITY PARAMETERS

The risk involved in the choice of a new variety can be attributed to two sources: uncertainty associated with environmental factors, and uncertainty about the performance of any particular variety, arising from the interaction between environmental factors and plant productivity. To quantify this risk, early approaches to the study of genotype-environment (GE) interactions by plant breeders were based on analyses of variance (ANOVA) of experimental results. The variance components were used to separate out the effects of genotypes, environments, and their interactions; in multisite, multiseason trials, the GE interaction was decomposed further into first order interactions, genotype × year and genotype × location, and a second order interaction, genotype × location × year. In general, the second order interactions tend to be much greater than the genotype × location or genotype × year effects (Moll and Stuber 1974; Ojomo and Adelana 1970). While ANOVA provides information about the magnitude of GE interactions, it ignores the stability of individual entries in the trial and does not help in the selection of a stable variety.

The Joint Regression Approach to Stability Analysis

As an alternative, Finlay and Wilkinson (1963) developed the joint regression approach for comparing varietal performance in several locations over several seasons to identify genotypic differences in adaptability. This is an extension of the work by Yates and Cochran (1938) and involves regressing yields on an environmental index. The mean yield of all varieties at each location in each season, the "site mean", is used as a measure of the environment's productivity. For each variety, a regression of individual yield on site mean yield is computed:

(1) $Y_{ij} = a_i + b_i \overline{Y}_{.j} + e_{ij}$

where Y_{ij} = the yield of the i^{th} variety at the j^{th} site, and $\overline{Y}_{.j}$ = the mean yield of all varieties at the j^{th} site in a given season. Using this procedure, Finlay and Wilkinson (1963) proposed the regression coefficient, b_i, as a quantitative measure of phenotypic stability: a) varieties with $b_i = 1$ have average stability. When this is associated with above average yields over all environments then the variety exhibits general adaptability (A in Figure 1); if the

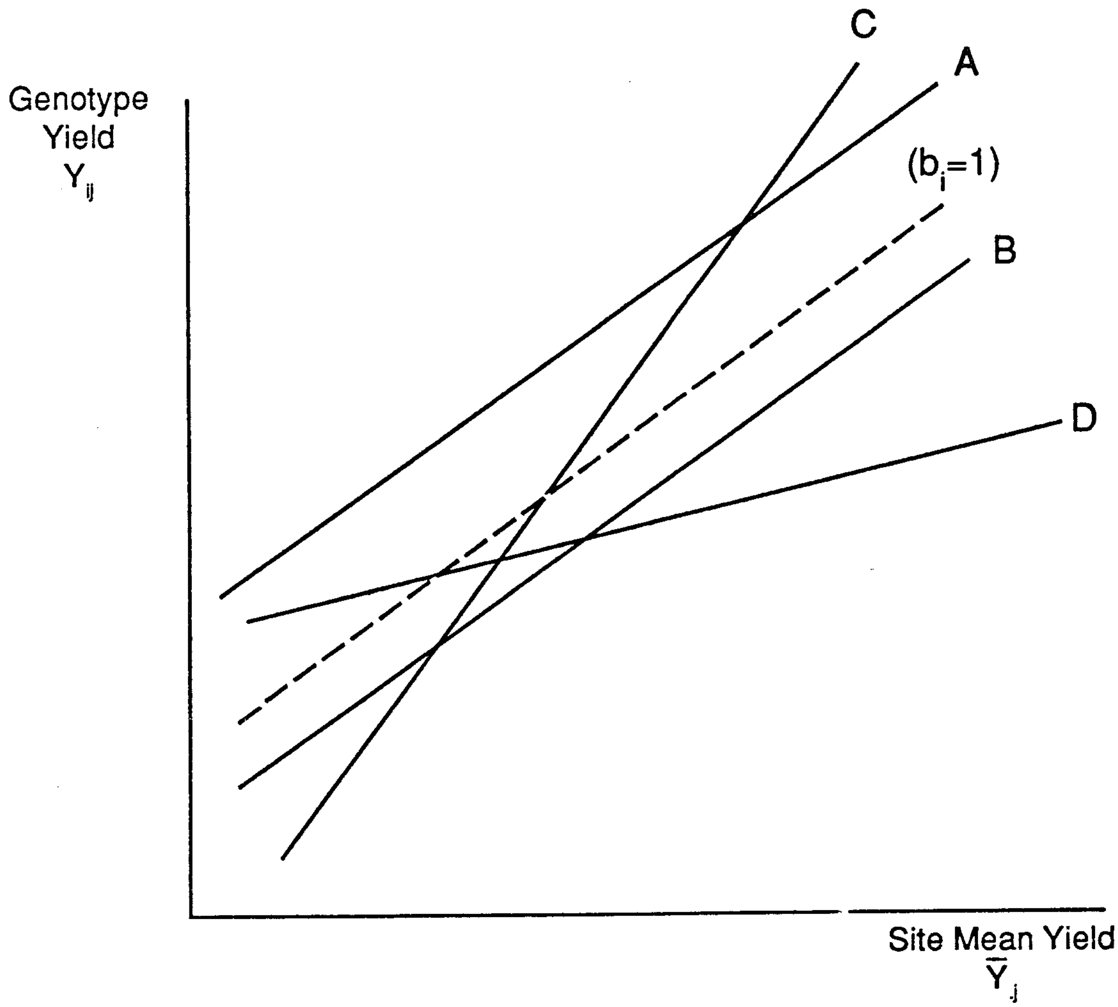

Figure 1. GE Interaction (adapted from Finlay and Wilkinson, 1963)

variety has below average yields in all environments then it is poorly adapted to all environments (B in Figure 1); b) varieties with $b_i > 1$ have below average stability; small changes in the environmental index produce large changes in yield. Such a variety (C in Figure 1), is specifically adapted to high yielding environments; and c) $b_i < 1$ indicates above average stability; the variety shows little change in yield despite large changes in the environmental index. Such a variety, being specifically adapted to low yielding environments (D in Figure 1), will have above average yields in poor environments but relatively low yields in better environments.

Finlay and Wilkinson (1963) use the term adaptability with reference to the yield of a variety relative to the average yield in any given environment, while stability refers to the degree of variation in yields of a variety across environments. A variety's mean yield over sites ($\overline{Y}_{i\cdot}$) is used as a comparative measure of its performance, and is plotted against b_i. Each variety is represented by a single point on the plot. One result of their study is that the variability (between varieties) in phenotypic stability (b_i) is inversely related to mean yield, giving rise to the triangular configuration of points in Figure 2.

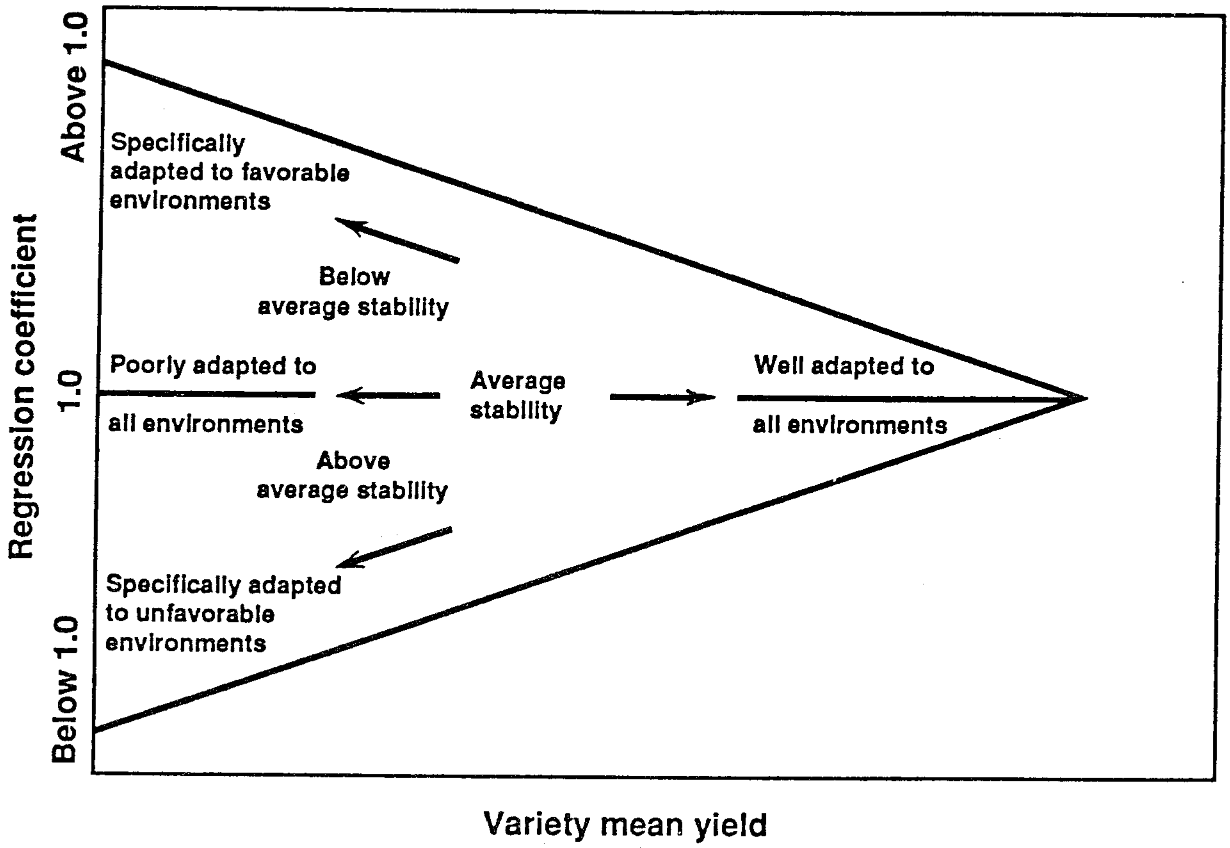

Figure 2. Generalized interpretation of varietal properties when stability parameters plotted against variety mean yields. (Source: Finlay and Wilkinson 1963).

Absolute phenotypic stability would be represented by a regression coefficient of zero. Finlay and Wilkinson define an ideal variety as the one with maximum phenotypic stability (b_i approaching 0), and maximum yield potential in the most favorable environment. However, they find that the varieties with high phenotypic stability all have low mean yields; they are "so stable, in fact, that they are unable to exploit high yielding environments" (Finlay and Wilkinson 1963, p.752). This implies that one must accept a compromise between stability and yield potential.

Eberhart and Russell (1966) use a similar model, regressing varietal yields on an environmental index. Although their environmental index was defined somewhat differently, the main contribution of the work was to suggest a second measure of stability, in addition to the slope coefficient. A variety that differs in its response from the majority of the varieties in a trial will have marked deviations around its regression line. These scientists suggest that a function of these deviations, S^2_{di}, be used as a second measure of stability, where S^2_{di} is the residual mean square corrected by a constant which is a measure of inexplicable environmental variation (the residual experimental error from the ANOVA table). They define an ideal variety as one with a high mean yield, a unit regression coefficient, and deviations, S^2_{di}, as small as possible.

This approach of regressing varietal yields on site means has intuitive appeal, is simple to apply and allows easy visual comparison of varieties. Consequently it has been used extensively by geneticists, plant breeders and other crop scientists as a means of ranking varieties by some measure of stability. However, the shortcomings of the approach have also been extensively discussed in the literature. Some authors take issue with the basic contention that GE interactions are linear functions of the environment (*e.g.* Knight 1970; Anderson 1974). Others have shown that regressions on site means do not satisfy fundamental statistical requirements, *e.g.* the site means are not statistically independent (Freeman and Perkins 1971), there exists bias in the estimates when estimated by ordinary least squares, and b_i and $S^2{}_{di}$ are not independent measures of stability (Hardwick and Wood 1972).

One further issue in this type of evaluation is the failure to distinguish between the temporal and the locational dimensions of GE interactions. Changes in a genotype's performance over time, within one location, can be attributed to fluctuations in the weather, while differences in performance between locations in a given year are a function of fixed locational factors, such as soil characteristics, as well as the prevailing weather conditions in each location.

To deal with this issue, Evenson *et al.* (1978) distinguish between the two sources of variability in performance. They define stability as low sensitivity to environmental changes over time, within a location. It is this environmental variation that is of importance to producers. Environmental variability across locations, on the other hand, while not being of concern to producers, has implications for crop improvement research when genotypes are being tested at multiple sites. Adaptability refers to the performance of a genotype with respect to changes in environmental factors across locations.

Evenson *et al.* (1978) also propose regressing genotypic yields on an environmental index, but use a two equation model to derive measures of stability and adaptability separately. They argue that mean yield of all genotypes within an environment is not a good measure of site potential; therefore, they regress yields of the i^{th} variety, at the j^{th} location in period t, on the average of the two highest yields, and a location dummy variable to remove systematic locational effects. The slope coefficient on the environmental index is then interpreted as a stability parameter. Similarly, the second equation contains a time period dummy variable and the slope coefficient is regarded as an estimate of an adaptability parameter.

Evenson *et al.* (1978) find that the two parameters are not closely related. They examine the relationship between stability and adaptability and the relative yielding ability of a variety by plotting the estimated stability and adaptability parameters against the average of the best two relative yields for each variety. They find that high levels of stability and adaptability do not appear to reduce the performance of varieties in the locations to which they are best adapted. However, it should be noted that they are dealing with data from international trials, where one would expect differences in locational environmental factors to be large. Finlay and Wilkinson (1963) used data from barley trials at a number of sites, all in South Australia, and Eberhart and Russell (1966) used data from trials all in north Iowa, which implies a far greater homogeneity of locational factors.

A final objection to methods that employ regression of varietal performance on environmental indices is that the results are both location and nursery specific. The regression coefficients are specific to the nursery (set of genotypes) tested; a genotype identified as being stable, on the basis of a low regression coefficient, is stable only relative to the other genotypes in the trial. Similarly, the deviations around the regression ($S^2{}_{di}$), proposed by Eberhart and Russell (1966) as a second measure of stability, are specific to the environments in which the trials were conducted.

Regression on Independent Environmental Variables

The use of a structural model of the relationship between genotype performance and environmental variables may eliminate the problem of nursery and locational specificity inherent in the previous approaches. In those models the slope coefficients are measures of a genotype's sensitivity to the totality of environmental factors. When critical environmental variables can be measured, they can be included as factors in each genotype's production function, and their effects quantitatively assessed. This approach, as used by plant physiologists, has in the past been based largely on results from greenhouse trials where the performance of a plant in response to controlled and measured factors can be easily recorded. Under field conditions, the necessary measurements of environmental factors have seldom been available. However, much of the more recent agronomic and agro-meteorological literature is concerned with the measurement and modelling of the relationship between plant performance and such environmental factors as daylength, temperature and soil moisture under field conditions.

Hardwick and Wood (1972) regress yields on environmental factors that are measured independently of plant performance. The regression coefficients in their model can be interpreted as stability or adaptability parameters. How the coefficients are interpreted will depend on the environmental variables included. Binswanger and Barah (1980) identify three types of plant independent variables: 'control variables' such as fertilizer, 'site variables' that vary across location, but not over time, such as soil characteristics, and 'weather variables', such as rainfall which vary across location and over time.

Focusing on the effect of weather variables, if a stable variety is defined as one that is insensitive to variations in weather, then the regression coefficients on the weather variables should be close to zero for a stable variety. In effect "we are looking for low explanatory power of the regression" (Binswanger and Barah 1980, p.15). This has implications for relevant hypothesis tests and the interpretation of measures of 'fit'. Other problems will involve errors arising from possible misspecification of the model or the omission of relevant variables. If an omitted variable is correlated with a variable included in the model then parameter estimates will be biased. An understanding of the relationship between environmental variables is needed before any conclusions can be drawn on their interactions with plant performance.

Some Final Observations

Despite the progress made in resolving the issues surrounding sample design and the specification of an appropriate structural model, much of the literature on genotype-environment interaction fails to discuss how, and by what criteria, the trade-off between yield and stability is to be made. One exception is the approach proposed by Binswanger and Barah (1980), results of which are reported in Barah *et al.* (1981). They identify stability as the converse of the degree of risk faced by the producer; a more stable variety is less 'risky'. Using multisite, multiseason trial data, the stability and adaptability components of the overall variance in yield of each genotype are estimated using the mean squares from the ANOVA table for each genotype. The estimated stability standard deviations are then plotted against mean yields. The genotypes can then be classified in terms of risk (stability) efficiency. A genotype is defined as risk efficient if no other genotype with the same mean yield has a lower standard deviation, or no other genotype with the same standard deviation has a higher mean yield. It is not possible to assess the trade-off between yield and standard deviation in the resulting efficient set without knowing something more about decision-makers' attitudes toward risk and returns. This mean-variance approach is one of a number of methods discussed in the following section on choice under uncertainty.

III - CHOICE UNDER UNCERTAINTY

In this section a variety of approaches to decision making under uncertainty are reviewed, before discussing their application to the problem of genotype selection under environmental variability.

Review of Expected Utility Maximization and Risk Aversion

In a risky situation, the decision maker must choose between alternative courses of action whose outcomes are determined by the state of an uncertain environment. Letting a_j = the j^{th} act or alternative course of action; s_i = the i^{th} state of the environment; $p_i = P(s_i)$ = the probability that s_i occurs; and x_{ij} = the outcome of a_j given that s_i occurs, the expected utility hypothesis is based on preferences that are consistent with the axioms of ordering, transitivity, continuity, and independence, for which there exists a utility function, $U(.)$, such that: a) if any risky action, a_1, is preferred to another, a_2, then $U(a_1) > U(a_2)$, and b) $U(a_j) = E_iU(x_{ij}) = \sum_i p_iU(x_{ij})$. The optimal act, a_j*, is that which maximizes expected utility (Anderson, Dillon and Hardaker 1977):

$$(2)\quad U(a_j^*) = \underset{j}{Max}\, U(a_j) = \underset{j}{Max}\, [\sum_i p_iU(x_{ij})].$$

The utility function is taken to be a single valued function of some measure of wealth, x. Several increasingly restrictive conditions may be imposed on the utility function, $U(x)$.[1]

First, it is assumed that individuals prefer more wealth to less; this implies a monotonically increasing utility function with marginal utility of wealth strictly positive, $U'(x) > 0$. Second, it is generally assumed that the utility function $U(x)$ exhibits decreasing marginal utility of wealth implying a concave function with $U''(x) < 0$; this is equivalent to assuming risk aversity.

A risk averse individual prefers a sure amount to taking a risk, *i.e.*, $U[E(x)] > E[U(x)]$, as long as the sure amount is not less than the minimum outcome from the gamble. This is demonstrated, for a simple lottery, in Figure 3. Suppose an individual is given the choice of playing a lottery that pays X_1 units of x with probability p_1 and X_2 units of x with probability $p_2 = 1 - p_1$. The expected outcome is $E(x) = \sum_i p_iX_i$. As can be seen from Figure 3, when the utility function is concave, the expected utility of the lottery, $E[U(x)]$, is less than the utility of the expected outcome:

$$(3)\quad E[U(x)] = \sum_i p_iU(X_i) < U[E(x)] = U[\sum_i p_iX_i].$$

The certainty equivalent, CE, is the amount, in units of x, such that $U(CE) = E[U(x)]$; it is the certain amount that will give the same utility as the lottery. Most risk averse individuals are willing to pay an insurance premium to avoid the uncertainty involved in the lottery. Pratt's risk premium, $\Pi(x)$, is the difference between the certainty equivalent and the expected outcome of the lottery; for risk averse individuals, $\Pi(x) > 0$.

---

[1] In many economic applications, wealth might be measured in terms of annual income or some measure of long-term asset position. In applying these decision models to genotype selection, if there are no differences in product quality, then one can ignore prices and assume that yield is the appropriate measure of wealth.

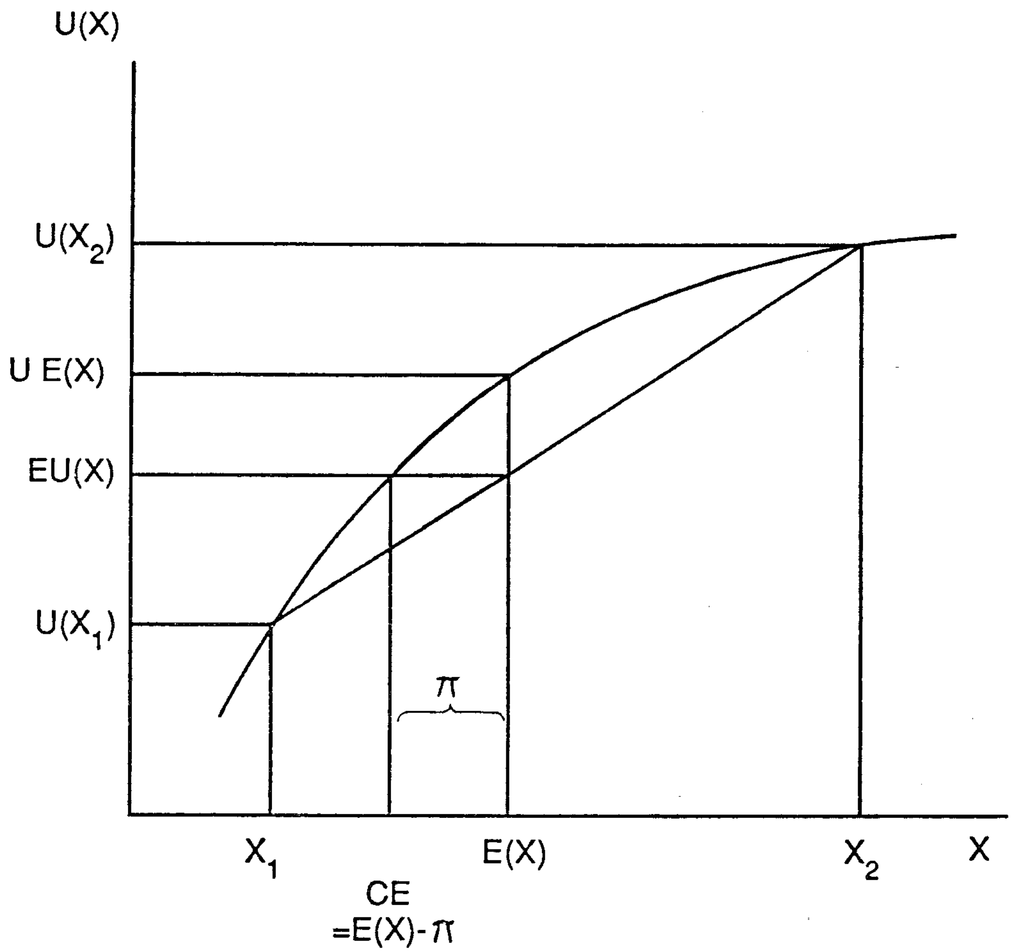

Figure 3. Illustration of the concepts of risk aversity, certainty equivalence, and the risk premium, π

The single valued utility function $U(x)$ is not a unique representation of preferences; any positive monotonic transformation of a utility function leaves the ranking of certain outcomes unchanged. The same does not hold for the ranking, in terms of expected utility, of uncertain outcomes. Expected utility rankings are invariant under any positive linear transformation of the form: $V(x) = a + bU(x)$, $b > 0$ (Henderson and Quandt 1980). While the sign of the second derivative, $U''(x)$, provides an indication of an individual's attitudes toward risk, its magnitude is no indicator of the degree of risk aversity because $U''(x)$ is not invariant under such linear transformations. The degree of risk aversity is, however, uniquely measured by the Arrow-Pratt absolute risk aversion function:

(4) $r_A(x) = - U''(x)/U'(x)$,

Values of $r_A(x)$ are local measures of the degree of concavity or convexity of a utility function and are unique measures of preferences; it can be shown that $r_A(x)$ is unchanged by any positive linear transformation of $U(x)$:

if $V(x) = a + bU(x)$, $b > 0$,

$V'(x) = bU'(x)$, $V''(x) = bU''(x)$, and

$r_A(x) = -bU''(x)/bU'(x) = -U''(x)/U'(x)$.

Relative risk aversion is defined as:

(5) $r_R(x) = -xU''(x)/U'(x) = xr_A(x)$

Arrow (1965) suggests that utility functions for risk averse individuals should display decreasing absolute risk aversion (DARA), *i.e.*, the degree of risk aversion decreases as wealth increases, but increasing relative risk aversion (Boisvert 1972). The assumption of DARA constrains the first derivative of the absolute risk aversion function to be negative:

(6) $r_A'(x) = \{U''(x)^2 - U'(x)U'''(x)\}/U'(x)^2 < 0$

Given the conditions for risk aversity: $U'(x) > 0$ and $U''(x) < 0$, this implies a further condition on the utility function; a positive third derivative is a necessary (but not sufficient) condition for DARA. A necessary and sufficient condition for DARA is $U'(x)U'''(x) > U''(x)^2$.

Bearing these conditions in mind, the discussion now turns to the application of expected utility (EU) maximization and the development of other selection criteria.

Empirical application of EU maximization involves identifying the alternative actions, a_i, specifying all possible states of nature along with their probabilities of occurrence, identifying the possible outcomes of each action, x_{ij}, given the alternative states of the environment, deriving a measure of the utility of each outcome, $U(x_{ij})$, and hence determining the utility of each action $E_i[U(x_{ij})]$. The action maximizing EU will be selected. This 'decision theoretic' approach to EU maximization assumes that preferences are completely known, *i.e.*, that a single valued utility function can be determined, and that probabilities can be specified.

When preferences are known and can be precisely formulated, the decision theoretic approach to maximizing expected utility gives a unique and complete ordering of actions, but in applied problems preferences are rarely known, are difficult to measure, and are not unique across decision makers. In many cases, however, individual decision makers' preferences may not be required; for instance, when dealing with policy questions, one is more interested in specifying how a group of individuals with similar preferences might respond or, in the case of plant breeders, one may be acting as an agent representing the preferences of a group of individuals. Under these circumstances other ordering criteria can be specified. Such criteria, in the absence of complete information on preferences, provide a partial ordering of alternatives by identifying two subsets: those that are 'risk efficient', for which no clear preference can be determined without further information on preferences, and those that clearly would not be preferred by any individual in the group (Boisvert 1985).

That is, given a set of conditions, or restrictions, placed on the set of utility functions of a group of individuals, then prospect A is preferred to prospect B, in terms of expected utility, if $E_AU \geq E_BU$ for every utility function in the defined set. Such a criterion is a sufficient condition for expected utility maximization. The efficiency criterion is an optimal criterion if it is both a necessary and a sufficient condition for expected utility maximization. An optimal efficiency criterion minimizes the efficient set of choices by discarding those that are inefficient.[2] Any further reductions in the efficient set require further restrictions on the admissible set of utility functions. The more restrictive the assumptions about preferences, the smaller the efficient set. In short, an efficiency criterion is a set of necessary and sufficient conditions on the probability distributions of outcomes for

[2] Prospects are inefficient in the sense that they would never be preferred by an expected utility maximizer in the group of decision makers defined by the restrictions on the utility function.

one action to be preferred to the other, in terms of expected utility, by all individuals in a particular group defined by conditions placed on the utility function.

Efficiency Criteria 1: EV Analysis

Perhaps the simplest and most widely used approach to efficiency analysis is the expected return-variance (EV) approach. This approach equates risk with variance (or the standard deviation) and the EV criterion can be stated as: if *A* and *B* are two uncertain actions, and $\mu_A \geq \mu_B$ and $\sigma^2_A \leq \sigma^2_B$, with at least one strict inequality, then *A* is preferred to *B*. Any choice between alternatives in the risk efficient set involves a trade-off between mean and variance and requires knowledge of the decision maker's preferences.

Despite the widespread application of the EV criterion in the economics and financial literature, there are a number of important objections to its use. First, the EV approach is consistent with the EU hypothesis when utility can be specified as a function of the mean and variance only. Using a Taylor approximation to the unknown true utility function, the utility of a risky prospect, $f(x)$, can be expressed as a function of its mean and higher moments about its mean (Anderson *et al.* 1977, p.92). The mean and variance completely specify the normal, and log-normal, distributions; odd moments about the mean are equal to zero, and even moments about the mean are functions of the variance. Accordingly, when returns are normally distributed, whatever the form of the utility function, it may be specified in terms of the mean and variance through the Taylor series expansion around the mean (Anderson *et al.* 1977, pp. 192-193). Alternatively, regardless of the form of the distributions of outcomes, if a quadratic utility function is assumed, the EV criterion is also consistent with expected utility maximization (Boisvert 1972). However, both these assumptions that make the EV criterion consistent with the expected utility hypothesis have undesirable properties. The assumption of normally distributed outcomes is empirically unrealistic in many cases. Day (1965) has demonstrated that the distributions of yields may deviate substantially from normality. The quadratic utility function has a third derivative of $U''' = 0$; it follows that $r_A'(x) = U''^2/U'^2 > 0$, *i.e.*, the quadratic utility function exhibits increasing absolute risk aversion. This is a theoretically unacceptable restriction to place on preferences.

Second, from a more practical point of view, EV analysis identifies risk with variance which means that extreme gains, as well as extreme losses, are considered undesirable. There may be cases where an increase in variance is not undesirable, for instance, if it is accompanied by an upward shift in the location of the distribution. One can envisage a situation where one distribution has both a higher mean and a higher variance than a second distribution, and with all its outcomes lying to the right of those in the second distribution. Under EV analysis, no clear preference can be determined, while clearly the first distribution would be preferred by any individual who prefers more to less.

Efficiency Criteria 2: Stochastic Dominance

The development of the theory of stochastic dominance has provided alternative efficiency criteria. If the value of the cumulative distribution function (CDF) of a preferred choice never exceeds that of an inferior choice, then the preferred choice is stochastically larger than, or stochastically dominates, the inferior choice. The stochastic dominance criteria provide a means of selecting alternatives that are optimal, according to expected utility maximization, for a specified set of utility functions.

Initially three such criteria were developed, requiring increasingly restrictive assumptions about preferences. These are reviewed in some detail before proceeding to discuss Meyer's (1977) "generalized" dominance criteria, commonly referred to as stochastic dominance with respect to a function.

Ordinary Stochastic Dominance

It is useful at this point to specify the notation that is used in the following sections. Define two risky prospects, F and G, with continuous outcomes lying in the range $[a,b]$, having CDFs $F(x)$ and $G(x)$, where $F(x) = P[X \leq x] = \int_a^x f(t)dt$, and $dF(x) = f(x)dx$, where $f(x)$ is the probability density function, and $F(a) = 0$ and $F(b) = 1$.

Define $F_2(x) = \int_a^x F(t)dt$, and $F_3(x) = \int_a^x F_2(t)dt$. The expected value of a risky prospect F is: $E_F(x) = \int_a^b xdF(x)$. The expected utility of a risky prospect F is: $E_FU = \int_a^b U(x)dF(x)$.

Statistically, F stochastically dominates G, or is stochastically larger than G, if $F(x) \leq G(x)$ for all x, with at least one strict inequality. In its economic use, stochastic dominance is defined as: F dominates G by the i^{th} degree stochastic dominance if, and only if, $E_FU > E_GU$ for all $U \in U_i$, where U_i is a specified class of admissible utility functions.

*First Degree Stochastic Dominance (*FSD*)* was developed by Hadar and Russell (1969, 1971) and Hanoch and Levy (1969). It imposes mild restrictions on the utility function; preferences are restricted to the set of utility functions, U_1, that are monotonically increasing: $U_1 = \{U: U \text{ continuous}, U'>0\}$; it follows that $-\infty \leq r_A(x) \leq \infty$. The ordering rule for FSD is: F dominates G by FSD if, and only if, $F(x) \leq G(x)$ for all x in $[a,b]$ with a strict inequality for at least one value of x.

The proof of sufficiency involves showing that if $F(x) \leq G(x)$, or $F(x)-G(x) \leq 0$, then $E_FU > E_GU$, or $E_FU-E_GU > 0$. Necessity can be proved by contradiction. This involves showing that if the conclusion above does not hold, *i.e.*, if $F(x) - G(x) > 0$, then there exists a utility function $U \in U_1$, such that $E_FU - E_GU < 0$. Such proofs are given by Hadar and Russell (1969; 1971) and Hanoch and Levy (1969).

The rule for FSD can be interpreted as F dominates G if, for every value of x, $P[X < x]$ is not larger for F than it is for G. Graphically, this means $F(x)$ may never lie to the left of $G(x)$. If the distributions cross, then no conclusion regarding FSD can be drawn.

A corollary to the rule (proved by Hadar and Russell, 1971) is that if F dominates G by FSD, then all the odd moments of $F(x)$ are greater than those of $G(x)$; when x is in the range $[0, \infty]$, then all moments of $F(x)$ exceed those of $G(x)$. The first odd moment is the mean, therefore if F dominates G, $\mu_F > \mu_G$. The reverse does not necessarily hold, however; one cannot identify first degree stochastically dominant distributions simply by comparing their means.

Second Degree Stochastic Dominance (SSD) was developed independently by Hanoch and Levy (1969) and Hadar and Russell (1969). It assumes a further restriction on the utility function, that of decreasing marginal utility. $U_2 = \{U : U \text{ continuous}, U'>0, U''<0 \}$

identifies the class of all strictly concave functions and represents all risk averse individuals. This assumption restricts the absolute risk aversion function, $r_A(x)$, to the range $[0, \infty]$.

The ordering rule for SSD is: F dominates G by SSD if, and only if, $F_2(x) \leq G_2(x)$ for all x in $[a,b]$, with a strict inequality for at least one value of x. The proofs of sufficiency and necessity can be found in Hadar and Russell (1969).

Graphically, SSD is interpreted as F is preferred to G, by all decision makers who are risk averse if, and only if, the area under $F(x)$ is less than that under $G(x)$. When the CDFs cross, the area between $F(x)$ and $G(x)$ when $F(x)$ lies above $G(x)$ must be less than the area between them when $F(x)$ lies below $G(x)$.

Two corollaries are proved by Hadar and Russell (1971): a) if F dominates G by SSD, then $\mu_F \geq \mu_G$; and b) if $\mu_F = \mu_G$ and F dominates G by SSD, then $\sigma^2_F < \sigma^2_G$. Fishburn and Vickson (1978) show that when the CDFs cross only once, then F dominates G by SSD if, and only if, $\mu_F > \mu_G$.

It should be noted at this point that the mean-variance (EV) criterion is a special case of SSD. When the utility function is assumed to be quadratic, or when F and G are assumed to be normally distributed, then the SSD efficient set is the same as the EV efficient set.

*Third Degree Stochastic Dominance (*TSD*)* was developed by Whitmore (1970). A third restriction on $U(x)$, a positive third derivative, is imposed such that $U_3 = \{U\colon U$ continuous, $U'>0,\ U''<0,\ U'''>0\}$. Recall that, as was shown earlier, this is a necessary, but not sufficient, condition for decreasing absolute risk aversion (DARA); it is possible to have $U'''>0$ and not have DARA. It follows that U_3 can contain utility functions exhibiting constant and increasing, as well as decreasing, absolute risk aversion. The restrictions on U_3 are therefore not as restrictive as imposing DARA.

The ordering rule for TSD is: F dominates G by TSD if, and only if, a) $F_3(x) \leq G_3(x)$, for all x in $[a,b]$, with a strict inequality for at least one value of x, and b) $F_2(b) \leq G_2(b)$, *i.e.* if $F_3(x)$ never exceeds $G_3(x)$ and the total area under $F(x)$ is less than the total area under $G(x)$.

The proofs of necessary and sufficient conditions are given by Whitmore (1970). He also presents a corollary of his theorem: if $\mu_F = \mu_G$ and F dominates G by TSD, then $\sigma^2_F \leq \sigma^2_G$ (inequality is not strict).

Any further reductions in the efficient set will require the imposition of further restrictions on the utility function, but there is no theoretical justification for such restrictions. If there appears to be a theoretical foundation for assuming DARA, then, defining $U_d = \{U : U \in U_2,\ r_A'(x) < 0\}$, since U_d is contained in U_3 we can expect a dominance criterion for U_d (DSD) to be stronger than TSD. Fishburn and Vickson (1978) find that DSD and TSD are equivalent when $\mu_F = \mu_G$, but that with unequal means DSD is stronger. At the time of their writing the DSD criterion was limited to discrete distributions; Fishburn and Vickson could not give general necessary and sufficient conditions for DSD.

Stochastic Dominance With Respect to a Function

Ordinary stochastic dominance has some shortcomings. First, it may not be discriminating enough and a decision maker will be left with a large efficient set of choices. Any further reduction in the size of the set would require theoretically unacceptable restrictions on preferences. Second, the existing restrictions on preferences imposed by the SD criteria may be difficult to support theoretically, or may not conform to empirical findings. For instance, King and Robison (1981a), in measuring decision-makers' preferences, found that most individuals exhibited increasing risk aversion over lower income levels.

There is, therefore, a need for criteria that offer greater flexibility and discriminating power. Ordinary SD criteria describe classes of admissible preferences by assuming restrictions on the form of the utility function. Stochastic dominance with respect to a function (SDWRF), developed by Meyer (1977a; 1977b), is a criterion which orders risky actions for a particular group of decision makers defined by placing assumed, or measured, restrictions on the upper and lower bounds of their absolute risk aversion function: $r_A(x) = -U''(x)/U'(x)$. The interval can be as wide or as narrow as desired.

Meyer defines $U(r_1(x), r_2(x))$ as a group of agents with expected utility functions satisfying $r_1(x) \leq -U''(x)/U'(x) \leq r_2(x)$, where $r(x)$ is the absolute risk aversion function, $r_A(x)$. Meyer's criterion provides, for any two functions $r_1(x)$ and $r_2(x)$, the necessary and sufficient conditions for one CDF to be preferred, or indifferent, to another by all agents in the class $U[r_1(x), r_2(x)]$. In this respect, SDWRF is a generalized version of ordinary stochastic dominance: if $r_1(x) = -\infty$ and $r_2(x) = \infty$ then SDWRF will identify the same efficient set as FSD; if $r_1(x) = 0$ and $r_2(x) = \infty$ then SDWRF will identify the same efficient set as SSD.

Meyer's more general approach, by allowing the specification of explicit bounds on risk aversion, allows the ranking of distributions that could not be ranked by ordinary stochastic dominance. In addition, by varying the values of $r_1(x)$ and $r_2(x)$, the effects, on the choice of action, of changes in the degree of risk aversion can be examined.

Discussion

The decision theoretic approach to ordering uncertain actions has limited applicability due to the difficulties involved in eliciting accurate information on preferences, particularly in the form of a single valued utility function. The EV approach is consistent with the expected utility hypothesis only under certain conditions. The stochastic dominance approach is based on Bernoullian decision theory and expected utility maximization, but does not require explicit knowledge of preferences; it does, however, require specification of the probability distributions of the outcomes of each action. This has been the cause of some reservations regarding the empirical application of stochastic dominance criteria.

A necessary condition for FSD and SSD, is that the lower bound of a dominated distribution cannot be less than that of an unpreferred distribution, or that a dominant distribution may not have a greater probability of the worst possible outcome than an unpreferred distribution. This requirement focuses attention on the lower, left-hand, tails of the cumulative distribution functions. Many researchers are concerned about the emphasis this places on the estimation of the lower, extreme values and their associated probabilities. If the lower tails of two CDFs are very similar, then the results of a stochastic dominance ordering will be very sensitive to any measurement errors. However, it can be argued that it is precisely this area of the distribution that producers are concerned about, *i.e.*, the risk of falling into the lower tail, or the risk of some disastrous outcome occurring, and any approach to decision making that is to take account of this risk must, implicitly, deal with the same problem. Anderson *et al.* (1977) argue that we must either accept the best that we can do and try to estimate lower tails as accurately as possible, or be forced to declare efficiency analysis infeasible.

An associated problem is the lack of statistical tests for stochastic dominance results; we are unable to test whether the differences between CDFs, identified by stochastic dominance criteria, are significantly different.

Given the null hypothesis, H_0: $E_FU = E_GU$, the alternative hypothesis, H_A, is $E_FU < E_GU$ or $E_FU > E_GU$. A Type I error occurs when H_0 is rejected (H_A accepted) when H_0 is true; a Type II error when H_0 is accepted when it is not true. Inaccurate predictions of orderings represent a Type I error, while the inability to order actions represents a Type II error. Cochran, Lodwick and Robison (1982) propose that the measurement errors associated with estimating a single valued utility function in the decision theoretic approach to EU maximization will lead to large Type I errors. The stochastic dominance approach while reducing the probability of Type I errors, will do so at the expense of increasing the probability of Type II errors.

Pope and Ziemer (1984) have gone some way toward testing the power of the stochastic dominance approach. Studies in stochastic efficiency differ in the way that the distributions are estimated. One alternative is to assume a parametric family of probability distributions and use sample data to estimate the underlying population parameters; the other, a non-parametric approach, is to use empirically estimated probability distributions. Pope and Ziemer compared the two approaches, comparing the parametric, or "plug-in" method, using appropriate Maximum Likelihood parameter estimates, for the normal, log-normal and gamma distributions, with the estimated empirical distribution function. They identify three possible conclusions that can be drawn from the comparison of the estimated CDFs, $\hat{F}$ and $\hat{G}$, either: A: neither distribution dominates, B: $\hat{G}$ dominates $\hat{F}$, or C: $\hat{F}$ dominates $\hat{G}$. Suppose that the correct (true) ranking in the population is that the true population distribution F dominates the true population distribution G. In comparing the estimated distributions, the occurrence of A or B results in an error regarding the conclusion about the dominance in the population. Pope and Ziemer (1984) argue that a Type B error (concluding that G dominates F when in fact F dominates G) is more serious than a Type A error (concluding that neither distribution dominates when in fact F dominates G), as it will result in a strictly erroneous, rather than a simply incomplete, ordering. The alternative estimators were compared in a Monte Carlo experiment using a number of parameter values and sample sizes, and their performance evaluated according to the relative percentages of correct rankings and the probability of Type A and Type B errors. They found that there was no noticeable difference between estimation methods, in the probability of Type A and Type B errors, and both declined as sample size increased. In terms of correct rankings, the empirical distribution compared favorably with the Maximum Likelihood methods, especially with small samples and regardless of the underlying parent distribution. They conclude that for most applied problems in agricultural risk analysis, the empirical approach is probably the best estimation method.

Efficiency Criteria 3: Mean Gini Analysis

The popularity of the EV approach to choice under uncertainty is attributable to its ease of application; one need only calculate and compare means and variances. An alternative approach, developed by Yitzhaki (1982), is based on a function of Gini's mean absolute difference. The approach has the convenience of the EV approach but does not equate risk with variance, and the decision rules are shown to be necessary conditions for SSD.

Gini's mean absolute difference is the expected value of the absolute differences between all pairs of values of a random variable with distribution $F(x)$:

(7) $$\Delta = E[|x-y|] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x-y| dF(x)dF(y).$$

It is dependent on the spread of the values among themselves and not on deviations from some constant value such as the mean. Gini's coefficient of concentration is $G = \Delta/2\mu$ (Kendall and Stuart 1958).

Yitzhaki proposes that a necessary condition for a distribution F_1 to dominate another, F_2, by FSD and SSD is: $\mu_1 \geq \mu_2$ and $\mu_1 - \Gamma_1 \geq \mu_2 - \Gamma_2$, with at least one strict inequality, where Γ_i is defined as one half Gini's mean difference:

(8) $$\Gamma_i = 1/2 \int\int |x-y| dF_i(x)dF_i(y),$$

which can be written as:

(9) $$\Gamma_i = \int F_i(x)[1-F_i(x)]dx$$

where $\Gamma = \Delta/2 = \mu G$. Proof of this proposition is given in Yitzhaki (1982). The same conditions are also derived from Yitzhaki's development of an extended Gini inequality index.

Yitzhaki (1983) defines the absolute parametric Gini index of equality for a distribution F defined over the range $[a\geq 0, b]$ as:

(10) $$\delta(v) = \int_0^b [1-F(x)]^v dx, \; v \geq 0,$$

with the following properties:

a) $\delta(v)$ is a non-increasing function of v:

(11) $$\partial\delta(v)/\partial v = \int_0^b \ln[1-F(x)]\,[1-F(x)]^v dx \leq 0$$

b) when $v=0$, $\delta(v)=b$;

c) when $v=1$, $\delta(v) = \int_0^b [1-F(x)]dx = b - \int_0^b F(x)dx = \mu$;

d) when $v=2$, $\delta(v) = \int_0^b [1-F(x)]^2 dx = \int_0^b [1-2F(x)+F(x)^2]dx$

$$= \int_0^b [1-F(x)] - [1-F(x)]^2 \, dx = \mu - \Gamma_F \, .$$

The absolute parametric Gini index of inequality is defined as

(12) $$\mu - \delta(v) = \Gamma(v) = \int_0^b [1-F(x)] - [1-F(x)]^v dx \, .$$

$\Gamma(v)$ is a measure of the degree of inequality, or the degree of dispersion, among points in a given distribution; $\delta(v) = \mu - \Gamma(v)$, the index of equality takes account of the mean and degree of inequality.

Yitzhaki (1983) defines necessary conditions for ranking two distributions according to $\delta(v) = \mu - \Gamma(v)$. Under the assumption $U'(x) \geq 0$, a necessary condition for F_1 to dominate F_2, in terms of expected utility is:

$$\delta_1(v) \geq \delta_2(v) \text{ for all } v \geq 0.$$

Under the additional assumption of risk aversity, $U''(x) \leq 0$, the necessary condition becomes:

$$\delta_1(v) \geq \delta_2(v) \text{ for all } v \geq 1.$$

If the two distributions cross only once, then these are also sufficient conditions for second degree stochastic dominance.

It follows that, under the assumption of risk aversity, a necessary condition for F_1 to dominate F_2 is that $\mu_1 \geq \mu_2$ and $\mu_1 - \Gamma_1 \geq \mu_1 - \Gamma_2$, where $\Gamma_i = \Gamma_i(2)$. This will be referred to as the Mean-Gini (MG) criterion.

An alternative way to look at this criterion is that the most equal (least variable) distribution will be that which maximizes the Gini index of equality $\delta(v)$ in (10). This index can be regarded as a weighted integration of the area under the CDF. A risk averse decision maker will place more weight on values in the lower tail than in the upper tail. Changing the value of v affects the weights attached to the points on the distribution; increasing v will increase the weights attached to the lower tails and decrease those attached to the upper tails. The parameter v can be regarded as a measure of aversion to inequality: $0 \leq v < 1$ represents aversion to equality, $v = 1$ represents indifference, and $v > 1$ represents inequality aversion (Yitzhaki, 1983). $F(x)$ can be interpreted as a measure of the rank of any given observation x: as x approaches the upper value b, $F(x)$ approaches unity, as x approaches its lowest value a, $F(x)$ approaches zero. It can be shown that $\delta(v)$ is a decreasing function of rank, F, and hence equality will be maximized, with respect to F, for all v, when observations are concentrated at b (Yitzhaki, 1983).

The construction of the MG efficient set is simpler than for SSD; it requires only the calculation of means and Gini's mean differences. Since the MG criterion is a necessary condition for SSD *i.e.*, SSD implies MG, every SSD inefficient (dominated) distribution will be inefficient by MG. The reverse is not necessarily true. Every MG inefficient distribution is not necessarily inefficient by SSD; a distribution may be inefficient, under the MG criterion, even if it is contained in the SSD efficient set. In short, while all MG efficient distributions are SSD efficient (by necessity), some SSD efficient distributions may be MG inefficient and the MG efficient set may, therefore, be smaller than the SSD efficient set. In other words the MG criterion exhibits greater discriminatory power. In contrast to the EV approach, which requires a trade-off between mean and variance, the MG criterion allows a prospect with a larger mean and a greater degree of dispersion to be preferred, if its higher mean compensates for the higher degree of variability. This supports Hanoch's and Levy's assertion that an increase in variability is not necessarily undesirable if it is accompanied by a shift to the right in the location of the distribution. It is not known, however, what the implications of this greater discriminatory power are, in terms of the admissible set of risk averse decision makers. If the MG criterion is reducing the SSD efficient set, it is doing so by rejecting some choices that may be preferred by some risk averse individuals in the admissible set defined by $U_2 = \{ U: U'>0, U''<0 \}$. The implication is that the MG criterion applies to some subset of U_2.

Buccola and Subaei (1984) suggest a number of reasons and produce results that support the hypothesis that the MG criterion best represents the preferences of relatively weakly risk averse decision makers. In one of the few empirical applications of MG analysis so far reported in the literature, Buccola and Subaei compare SDWRF (with varying absolute risk aversion intervals) with EV and MG criteria. The MG efficient set is identical to the SDWRF efficient set when $0 \leq r_A(x) \leq 0.0015$. As the upper bound on $r_A(x)$ is increased to 0.0045 and then to infinity (SSD), hence admitting more strongly risk averse individuals into the admissible set, the SDWRF efficient sets increasingly diverge from the MG efficient set. Buccola and Subaei conclude that, for representing weakly risk averse individuals, the MG approach deserves close attention.

In another study, Bey and Howe (1984) examine the empirical performance of the MG criterion compared to the EV, mean-semivariance (ES), FSD, SSD, and TSD criteria. They compare the resulting efficient sets in terms of size, common membership and properties of the efficient portfolios. The MG efficient set was by far the smallest; the average MG efficient set consisted of only 19% of the average SSD efficient set. In all cases MG efficient portfolios were SSD efficient; in addition, in most cases, they were also a subset of the TSD efficient set. There was a strong tendency for the MG efficient set to contain mostly those portfolios with high returns and high variances. Bey and Howe conclude that the MG criterion is potentially useful if the admissible set of decision makers could be more accurately defined.

Although it is not possible to define the precise absolute risk aversion intervals represented by the MG efficient portfolios, it might also be argued that this is no great disadvantage. Efficiency criteria were, after all, developed in response to the problems involved in the precise specification of preferences. Furthermore, as Yitzhaki (1982) points out, the MG approach is similar to Baumol's (1963) expected-gain confidence-limit (EL) criterion where decision makers choose alternatives that maximize μ subject to a given level of $L = \mu - \phi\sigma > 0$. For $\phi\sigma = \Gamma$, the criteria are the same and are not unlike other safety-first criteria. As with the EL approach, a prospect of greater risk can be preferred over another, by the MG criterion, if the mean is large enough.

Exponential Utility-Moment Generating Function Approach

A number of possible approaches to decision making under uncertainty have been discussed, based on the hypothesis that a decision maker will prefer the choice that maximizes his expected utility. Each approach has its advantages and disadvantages. Direct EU maximization requires explicit specification of the utility function and the statistical distribution of outcomes. The EV approach is only consistent with expected utility maximization under certain conditions. Results from efficiency analysis may be inconclusive. The final approach to be discussed, the exponential utility-moment generating function (EUMGF) approach, leads to a complete ordering of uncertain choices, according to expected utility.

The EUMGF approach was developed by Hammond (1974) who observed that the negative exponential utility function yields a simple expression for expected utility in terms of the moment generating function (MGF) of the random variable. Yassour *et al.* (1981) applied Hammond's approach to the case of discrete choices among technologies with stochastic yields. Their approach is simplified slightly by assuming, in our case, that costs and product prices are the same for all varieties. Let T_i be the i^{th} technology (variety) with stochastic yields Y_i and fixed price P, then returns to the i^{th} technology, $R_i = PY_i$. Under

EU maximization, T_i will be preferred to T_j if $E[U(R_i)] > E[U(R_j)]$. The EUMGF approach assumes a negative exponential utility function:

(13) $U(R_i) = -exp(-rR_i) = -exp(-rPY_i)$,

where r is a measure of absolute risk aversion. The MGF of a random variable Y_i, distributed F_i, is:

(14) $M_i(t) = E[exp(tY_i)] = \int_{\infty}^{\infty} exp(ty)dF_i(y)$.

Expected utility can then be expressed as:

(15) $E[U(R_i)] = -E[exp(-rPY_i)] = -M_i(-rP)$.

The certainty equivalent *(CE)* as defined in equation (3) is a sure return whose utility is equal to the expected utility of a risky prospect: $U(CE) = E[U(X)]$. It follows that:

(16) $U(CE) = -exp(-rCE) = E[U(R_i)] = -M_i(-rP)$, and

(17) $CE = -(1/r)ln[M_i(-rP)]$.

Maximizing expected utility is equivalent to maximizing the certainty equivalents of each technology. The technologies can then be ranked by their *CE*'s, under alternative specifications of the measure of absolute risk aversion, r_A. Yassour *et al.* (1981) derive ordering rules for a number of parametric distributions: Normal, Gamma, Chi-square, Exponential and Poisson. In the case of the Normal, EUMGF reduces to the EV criterion. In the other cases, the parameters of the MGF appearing in the decision rules can be expressed as functions of the means and variances of the random variable. Collender and Zilberman (1985) extended the approach to the problem of continuous choice and multivariate distributions, and derived rules for the optimal allocation of land among a number of alternative crops.

The EUMGF approach has many desirable properties: it leads to a complete ordering according to EU maximization, and ordering rules, based on means and variances only, can be derived for any parametric distribution with a finite MGF. Varying the parameter r_A allows actions to be ranked under varying degrees of risk aversion.

The use of the EUMGF approach also has its shortcomings. First, it depends critically on whether the assumption of a negative exponential utility function is an acceptable representation of preferences. The negative exponential utility function exhibits constant absolute risk aversion; as discussed earlier in this chapter, it has been argued that most individuals exhibit decreasing absolute risk aversion.

The second problem is that of selecting, under uncertainty about the exact distribution of returns, the parametric distribution (or MGF) that best represents the empirical data. Collender and Zilberman discuss various criteria for the selection of a distribution. However, a recent development of the approach, by Collender and Chalfant (1986), circumvents this problem. Their approach retains the assumption of an exponential utility function but replaces the parametric MGF with a non-parametric estimator, the empirical MGF (EMGF).

Given a random sample of independent and identically distributed random variables X_i $(i = 1, \ldots N)$, the quantities $exp(tX_i)$ are also independent and identically distributed; the EMGF is the sample mean of these variables:

$$(18) \qquad \hat{M}_i(t) = N^{-1} \sum_{i=1}^{N} exp(tX_i).$$

The selection criteria becomes:

$$(19) \qquad \text{maximize } CE = -(1/r)\ ln[N^{-1} \sum_{i=1}^{N} exp(-rX_i)].$$

Quandt and Ramsey (1978) show that the EMGF is a minimum variance, unbiased estimator of the true, unknown MGF. This non-parametric approach allows complete ordering of uncertain choices without requiring parametric specification of the distribution functions as long as one is willing to accept the assumption of constant absolute risk aversion.

Application to Selection of Genotypes

A comparison of genotypes using the various decision criteria discussed above, requires a set of observations on the performance of the genotypes over a time span of sufficient length to provide an estimate of the probability distribution of yields. In the absence of such observations on genotype performance, the assessment and choice of improved genotypes has largely been based on data from multisite, multiseason trials.

In section *II* various approaches to the assessment of genotypes in terms of their interaction with the environment are discussed. The standard joint regression approach does not distinguish between the temporal and locational dimensions of yield variability, and the resulting 'stability' parameters are location and nursery specific. It is shown that any choice between genotypes would involve a trade-off between yield and 'stability'. To rank the genotypes according to their expected utility would require that preferences be measured in terms of means and the 'stability' parameters.

The approach reported by Barah *et al.* (1981) is an illustration of the use of EV analysis. The optimal genotype was found using the slope (ratio of mean to standard deviation) of the iso-utility curves as measured by Binswanger (1980). This approach has all the shortcomings discussed earlier in this section. It assumes either that yields are normally distributed or imposes a quadratic utility function, thereby equating risk with variance (in this case the stability component of yield variance). Binswanger and Barah (1980) show that the results are also nursery and location specific. Risk efficient genotypes are efficient only relative to others in the nursery and given the environments included in the multisite, multiseason trial.

The classic example of the application of stochastic dominance to the examination of how different genotypes perform under uncertain environments is that by Anderson (1974). Anderson uses data from the Sixth International Spring Wheat Nursery, in which 49 varieties were compared in trials at 60 locations covering 39 countries. For each variety, the sixty trial observations on yield make up a discrete sample probability density function (pdf); each observation is assumed to have equal probability of occurring. The sample pdf's therefore reflect the distribution of yields for each variety, given the environments in which the trials were conducted.

All approaches to the appraisal of new varieties, when based on data from multisite, multiseason trials suffer from the fact that their results are specific to the environments included in the trials. Their usefulness in identifying generally stable, or adaptable, varieties is recognized. The problem comes when one wishes to make recommendations for one particular area. Producers are concerned only with the variability in production over time. Whether a variety will be adopted by a group of producers will depend on the environmental conditions in which they operate, and on how the variety will perform given those conditions. Any recommendations must take account of the environmental variability, over time, in a given area.

As discussed in section *II*, the analysis of GE interactions using independent environmental variables is neither location nor nursery specific. However, instead of interpreting the regression coefficients as stability or adaptability parameters, it is proposed that, if the trials include a wide enough range of environments, the model can be used to predict yields for any given environment for which the relevant factors have been measured. If the probability distributions of those factors can be specified for a given location, then the distribution of yields can be derived. The decision criteria discussed in this section can then be applied, and the results will no longer be specific to the trial environments, but will be specific to the location of interest.

Much of the variability in crop production in semi-arid areas can be attributed to variable climatic factors, particularly available moisture. It is proposed that one way to stabilize production in such areas is through the introduction of drought tolerant varieties. The identification of potential varieties requires trials that include a sufficient range of rainfall regimes. Such an experiment has been conducted by Groundnut Physiologists at the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), in India.

ICRISAT examined the effect of varying degrees of water stress on the yields of 22 genotypes of groundnut. Irrigation was used to simulate a whole range of drought conditions varying in their timing, duration and intensity. The experimental results have been used to measure the relationship between yield and different patterns of water stress. However, recommendations for a particular area must take account of the incidence of drought conditions in that area, and consider farmers' preferences regarding any trade-off between yield and drought tolerance (stability).

Historical data on the relevant meteorological variables were used in the specification of empirical probability distributions of drought conditions for three groundnut producing areas of India, Hyderabad, Anantapur and Gujarat. Combining these distributions with the experimental results makes it possible to simulate empirical yield distributions and, in turn, evaluate the 22 genotypes in these three locations according to the various decision criteria discussed above. Before proceeding with a description of the experimental data and procedures for estimating response functions, some background on groundnut production is provided.

IV - GROUNDNUTS IN INDIA: BACKGROUND AND DATA SOURCES

This case study addresses the question: which genotypes should be recommended for three groundnut producing regions of India, where recommendations are based on yield re sponse to water stress? The analysis, which is reported in greater detail by Bailey (1988), is conducted in four steps: a) historical meteorological data in the regions are used to derive empirical distributions of water stress; b) experimental results are used to model the relationship between genotype performance and critical environmental factors, in this case the timing, duration and intensity of water stress; c) results from (a) and (b) are combined to derive yield distributions for each genotype in each region; and d) the yield distributions are used to evaluate genotypes for each region according to the alternative selection criteria discussed in section *III*.

This section contains a brief discussion of groundnut production and its importance to the Indian economy, and a description of the three selected sample locations. The design of the trial, conducted by ICRISAT, on the response of 22 genotypes to water stress, is then discussed. Subsequent sections deal with the measurement of water stress and the empirical distributions of water stress in each of three sample locations, the estimation of a response relationship between yield and water stress, and the generation of yield distributions.

Groundnut Production in India

Arachis hypogaea L. (Groundnut, Peanut) is a leguminous oil seed, distinguished from other such species by the fact that its fruits mature underground. It is adaptable to a wide range of soil and climatic conditions. Its major use worldwide is as a source of cooking oil, but it is also processed into high protein products, such as peanut butter. Whole nuts are used for confectionary purposes, and whole roasted nuts are a delicacy worldwide. The oilcake remaining after oil extraction is also a nutritious animal feed; the groundnut haulms (or hay) remaining after harvest may be used locally as livestock fodder.

Commercialization of groundnuts has made it an important cash crop in many countries. Approximately 70% of the world's production is grown in the rainfed areas of the semi-arid tropics (ICRISAT 1983). India is the largest producer, accounting for over 30% of the world's production, followed by China and the United States (Table 1).

Average yields in India, Senegal, Sudan and Nigeria, are low compared with those of China and, particularly, the United States. Low yields can be attributed to two major factors: unreliable rainfall, and biological constraints, such as pests and disease. Groundnuts in India and Africa are grown primarily under rainfed conditions and therefore suffer the effects of unreliable, and frequently, insufficient rainfall.

The area planted to groundnuts in India increased steadily in the 1950's and early 1960's (Central Statistical Organization, 1984) (Table 2). Over the period 1980-1984, it remained relatively stable, averaging 7.3 million ha., with a coefficient of variation of about 5%; over the same period total production ranged from approximately 5 million metric tons (M.T.) to 7.3 million M.T., with a much higher coefficient of variation of about 20%, illustrating the effect of variability over years in rainfall.

India is a leading exporter of groundnut cake, (Table 3), but it exports less than one percent of its production in the form of whole nuts. This is in sharp contrast to the United States which exports 18-21% of its production as whole nuts. Whole nuts are normally destined for the confectionary market and must meet strict requirements regarding size, color, flavor, etc.

Groundnuts produced in India are used predominantly for oil extraction; India exports no groundnut oil, retaining its entire production for domestic consumption. India produced 884,000 M.T. of vegetable oil in 1982-83, and 890,000 M.T. in 1983-84 (Central Statistical Office 1984). This figure includes oil from eight other oilseeds grown in India but groundnuts are by far the most important oilseed, making up over 40% of the area, and over half the total production, of edible oilseeds (Directorate of Economics and Statistics 1984).

Table 1. Area, Production and Average Yield of Groundnuts in Six Major Groundnut-Producing Nations

		Area	Production	Yield	% World Production
		('000 ha.)	('000 M.T.)	(kg./ha.)	
World	1974-76[a]	18,920	17,830	942	
	1979-82[b]	18,844	18,673	991	
	1984[a]	18,350	20,611	1,123	
India	1974-76	7,109	5,720	803	32
	1979-82	7,185	5,887	817	32
	1984	7,250	6,900	952	34
China	1974-76	1,909	2,247	1,177	13
	1979-82	2,385	3,594	1,502	19
	1984	2,442	4,900	2,007	24
Senegal	1974-76	1,267	1,228	969	7
	1979-82	1,032	691	674	4
	1984	873	682	782	3
Sudan	1974-76	795	849	1,068	5
	1979-82	971	823	848	4
	1984	800	420	525	2
Nigeria	1974-76	937	393	420	2
	1979-82	600	573	954	3
	1984	600	550	917	3
United States	1974-76	606	1,701	2,801	10
	1979-82	575	1,553	2,700	8
	1984	614	2,008	3,270	10

[a]FAO Production Yearbook 1984

[b]Directorate of Economics and Statistics 1984

The Three Sample Locations

In selecting sample locations for the study, our attention was focused on major groundnut producing regions. Table 4 presents the area, production and average yields of groundnuts in the ten top producing states in India in 1983-84. Between them, Gujarat and Andhra Pradesh produced just under half of India's output; 79% of Gujarat's production and 72% of the production from Andhra Pradesh are produced in the kharif, or rainy season.

The selection of specific sites within Andhra Pradesh and Gujarat was largely determined by the availability of meteorological data. Agriculturally significant droughts occur when a lack of precipitation leads to the depletion of soil moisture to the point where plants experience stress. A shortfall in expected rainfall is not, in itself, an indicator of drought; drought only begins when available soil moisture is exhausted. Thus, when considering

Table 2. Area, Production and Average Yield of Groundnuts in India

Year	Cropped Area	Area: Groundnut	Area: % of Cropped	Groundnut Production	Average Yield
	---- ('000 ha.) ----			('000 M.T.)	(kg./ha.)
1950-51	131,893	4,494	3	3,481	775
1955-56	147,311	5,133	4	3,862	752
1960-61	152,772	6,463	4	4,812	745
1965-66	155,276	7,698	5	4,263	554
1970-71	165,791	7,326	4	6,111	834
1975-76	170,995	7,222	4	6,755	935
1980-81	172,305	6,801	4	5,005	736
1981-82	174,764	7,429	4	7,223	972
1982-83	169,657	7,215	4	5,282	732
1983-84	173,324	7,640	4	7,284	953

Source: Central Statistical Organization 1985.

agricultural droughts, other meteorological variables, besides rainfall, must be considered, such as temperature, wind velocity, and evapotranspiration, as well as other factors, such as the moisture holding capacity of the soil and the stage of crop growth which will determine moisture use. Meteorological data for two sites, Hyderabad and Anantapur, in Andhra Pradesh, and for four sites in Gujarat, were provided by ICRISAT.

Hyderabad (18°N, 78°E) is the location of ICRISAT's experimental station. Daily rainfall data are available for the years 1901-1984 and other daily meteorological data, from ICRISAT's meteorological station, for 1974-1985.

Anantapur (15°N, 77°E) lies to the south of Hyderabad; Anantapur District is one of the leading producers of groundnuts in Andhra Pradesh. In 1982-83, 43% of the total cropped area in Anantapur District was planted to groundnuts, producing 236,040 M.T., 21% of the state's total production of groundnuts. Daily rainfall data are available for the years 1911-1984, while other daily meteorological data are available for one year only, 1985.

As shown in Table 4, Gujarat is the foremost groundnut producing state. While weekly rainfall data were available for four sites in Gujarat, Class A pan evaporation data (1983-1985) were available only for Anand (22°N, 73°E) in Kaira District. While Anand is not in the primary groundnut-producing area of Gujarat, it is considered reasonably representative of Gujarat's climate and it was decided to use the rainfall data for the years 1952-1983 from Anand rather than combine Anand's evaporation data with rainfall data from another site.

The three sites differ climatically. Figures 4, 5, and 6 show the differences in average weekly rainfall distributions and evaporative demand across the three sites. (See Bailey 1988 for more detailed meteorological data.)

Hyderabad has three distinct seasons. The rainy season (monsoon, kharif) usually begins in June and extends into early October. The average annual rainfall (1901-1984) is 780 mm. of which 85% falls in the kharif season; the peak rainfall months are July through

Table 3. Quantity and Value of Exports by India and Other Leading Net Exporters of Groundnut Products

	Quantity			Value		
	1982	1983	1984	1982	1983	1984
	- - - - - - (M.T.) - - - - - -			- - - - - - ($'000) - - - - - -		
Cake						
India	258,491 (37)[a]	236,054 (33)	319,413 (53)	37,800 (35)	32,000 (31)	45,000 (50)
Senegal	194,877 (28)	197,000 (28)	110,000 (18)	30,906 (28)	29,000 (28)	17,000 (19)
Sudan	72,458 (11)	83,308 (12)	36,000 (6)	13,111 (12)	11,516 (11)	6,300 (7)
Whole nuts (shelled)						
India	36,000 (5)	21,500 (3)	38,000 (5)	32,000 (6)	19,600 (4)	36,000 (7)
USA	200,037 (28)	222,399 (30)	263,970 (36)	171,971 (33)	173,971 (35)	200,642 (36)
China	107,570 (15)	154,302 (21)	143,100 (20)	76,051 (15)	91,002 (18)	102,000 (18)
Sudan	88,990 (12)	15,510 (2)	36,000 (5)	34,302 (7)	12,341 (3)	30,000 (5)
Oil						
Senegal	151,485 (34)	154,810 (31)	108,000 (31)	95,242 (33)	93,243 (31)	108,000 (31)
China	54,700 (12)	51,271 (10)	64,181 (19)	35,000 (12)	30,979 (10)	61,264 (18)
Brazil	77,623 (17)	56,962 (12)	26,455 (8)	44,985 (16)	27,467 (9)	24,943 (7)

Source: FAO Trade Yearbook 1984

[a]Figures in parentheses are % of world exports.

Table 4. Ten Leading Groundnut-Producing States in India, 1983-84

State	Cropped Area	Groundnut Area	% Cropped Area	Production	% All India	Average Yield
	- - - ('000 ha.) - - - -			('000 M.T.)		(kg./ha.)
Gujarat	10,695	2,150	20	1,905	26	886
Andhra Pradesh	12,281	1,629	13	1,695	23	1,041
Tamil Nadu	6,469	1,061	16	1,073	15	1,011
Maharashtra	20,270	812	4	834	11	1,027
Karnataka	10,660	843	8	747	10	886
Orissa	8,746	279	3	378	5	1355
Madhya Pradesh	21,402	345	2	247	3	716
Rajasthan	17,350	183	1	174	2	951
Uttar Pradesh	24,574	249	1	160	2	643
Punjab	6,763	58	1	42	1	724

Source: Central Statistical Organization 1985.

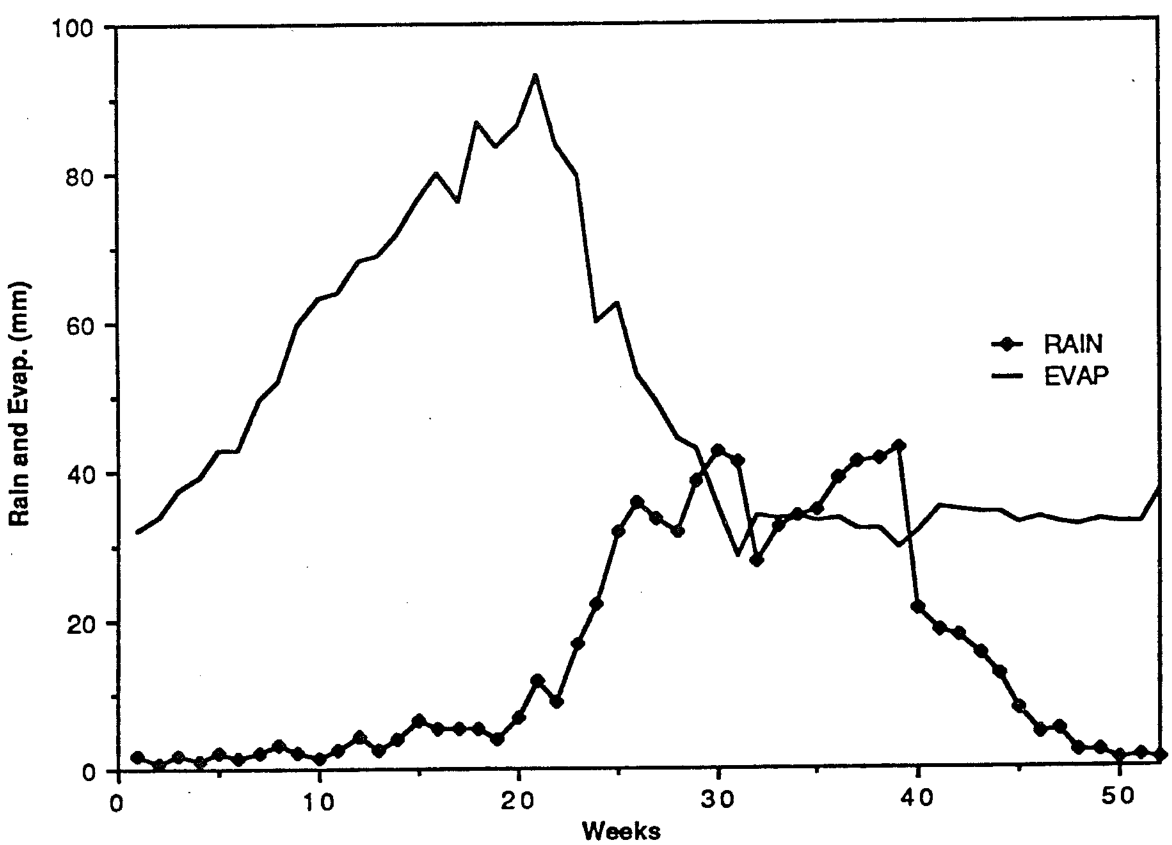

Figure 4. Mean weekly rainfall and evaporation, Hyderabad

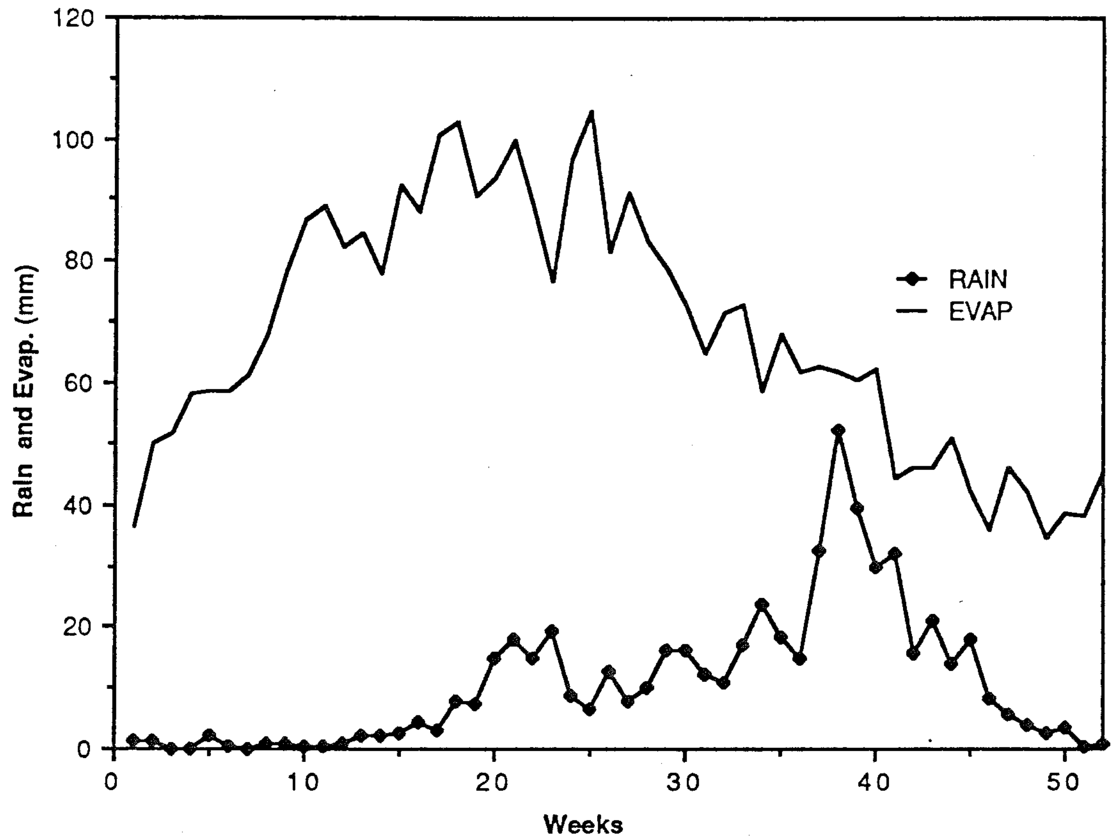

Figure 5. Mean weekly rainfall and evaporation, Anantapur

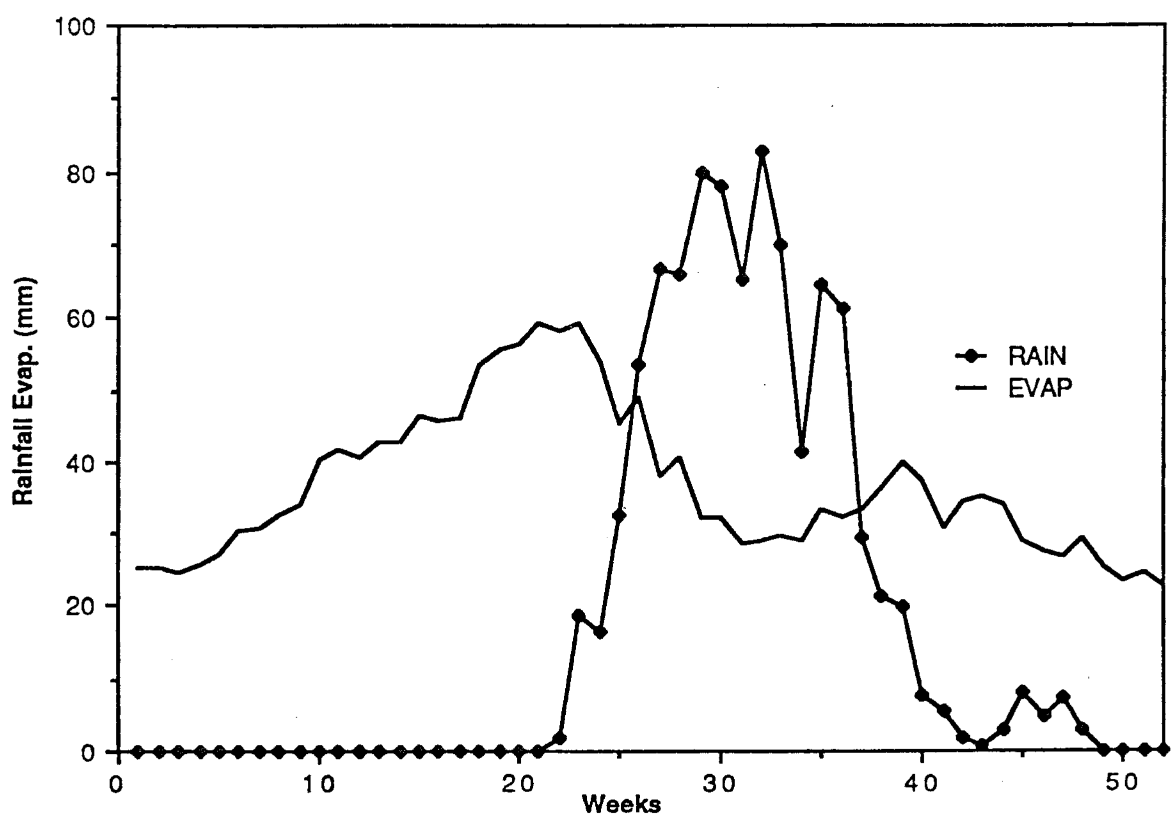

Figure 6. Mean weekly rainfall and evaporation, Gujarat

September when 63% of the average annual rainfall is received. Maximum daily temperatures decline from approximately 38°C in early June to 30°C in September-October, and evaporative demand (pan evaporation) falls from a daily average of about 12 mm. in early June to 4-5 mm. in October. The post rainy winter season (rabi) lasts from mid October to January and is dry and cool with shorter days. While maximum temperatures remain around 28°C-30°C throughout the season, minimum temperatures drop from 22°C in September-October to 12°C-15°C in December-January. The hot, dry summer season lasts from February until June when the rains begin. Maximum temperatures steadily rise, as does daily pan evaporation from 5 mm. in January to 12-13 mm. in May.

Anantapur, to the south of Hyderabad, has a longer rainy season (end of May until mid November), but receives only 560 mm. of rainfall annually. About 60% of the rainfall comes between August to October. Maximum temperatures remain around 32°C-33°C in the rainy season and this is reflected in higher daily pan evaporation which declines from 13-14 mm. at the end of May to 8-9 mm. in October and to 6 mm. in November. The lower rainfall and higher evaporative demand lead to a greater incidence of periodic droughts.

As can be seen from Figure 6, the rainy season in Anand, Gujarat is clearly defined and short; rainfall is restricted largely to the four months of June to September. However, average annual rainfall is 913 mm., 95% of which falls during those four months. Average daily pan evaporation is lower than in Hyderabad, declining from 8 mm. in June to 5 mm. in October. Crop production in this area is controlled by the short growing season, and any delay in planting is critical.

In short, while the three sample locations represent the foremost groundnut producing regions of India, they also highlight the agroclimatic differences in groundnut growing environments. Hyderabad has a relatively high assurance of adequate rainfall. Anantapur has a much lower rainfall, with a high and variable incidence of drought. Gujarat, with relatively high rainfall, has a high incidence of late season drought.

Experimental Data on Groundnuts

The identification of potential risk reducing genotypes for these three sample locations requires information on yield response to different patterns of drought. The severity of a drought, in terms of its effect on plant growth and productivity, depends on its intensity, or the degree of moisture deficiency, its duration, and its timing with respect to the stage of growth of the plant.

In 1982-83, ICRISAT conducted an experiment to study the effects of the three components of drought, timing, duration and intensity of water stress, on groundnut productivity, and the extent of genotypic variability. The experiment was conducted at ICRISAT's center near Hyderabad, in Andhra Pradesh. To allow simulation of drought patterns using irrigation, and without any interference from rainfall, the experiment was conducted in the late post rainy and summer season.[3]

The experiment was sown in early December, 1982, after a basal application of fertilizer, on a clay silt alfisol with available water holding capacity of approximately 100 mm. The crop was harvested in mid-April, 1983. Twenty-two groundnut genotypes of comparable maturity belonging to subspecies *fastigiata*, varieties *fastigiata* (valencia) and *vulgaris* (spanish) were included in the trial. The genotypes were selected to include lines which

3 The experimental design and preliminary results are described in detail by Nageswara Rao and Williams (1985); Nageswara Rao, Williams and Singh (1985); Williams et al. (1986). Only a general overview is presented here.

have been found to be tolerant, average or susceptible to drought in earlier drought screening trials at ICRISAT (Nageswara Rao and Williams 1985). They include established commercial cultivars and popular Indian cultivars as well as new accessions and advanced breeding lines. Details of the 22 genotypes and their GNO identifiers for this study are in Appendix A.

For the experiment, the land was divided into 12 blocks. Within each block, the land was prepared in beds and furrows. Each block contained eight beds, running the length of the block, of 1.2 m. width, divided by furrows 0.3 m. wide. Each genotype was sown in paired rows (one plot) of 12 m. length running across the eight beds. This arrangement was replicated three times within each block.

The crop was irrigated uniformly for 30 days after sowing (DAS), to ensure crop establishment and a fully charged soil profile. The drought treatments, designed to examine the effects of both single and multiple droughts, varying in timing and duration, were imposed thereafter. The drought patterns were developed in relation to four phenological phases in the groundnut growing cycle: a) seedling-flowering vegetative phase; b) pegging phase; c) podset phase; and d) podfilling to maturity. Twelve drought patterns were selected and assigned at random to the 12 blocks. Within each block (pattern) the genotypes had been assigned at random to the plots (paired rows) within each replicate. Line source irrigation (Hanks *et al.* 1976) was used to create eight intensities of drought within each pattern, by varying the irrigation amount systematically across the eight beds.

Table 5 provides details of each of the 12 patterns. Periods of stress are created by line source (LS) irrigation; release of water stress is represented by a uniform (U) irrigation. The 12 patterns include single periods of stress over each of the four phenological phases, plus various combinations of multiple droughts.

Table 5. Timing and Duration of Drought Patterns in the Experiment

Irrigation			Patterns											
Interval	DAS	Days	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
1	0	15	U	U	U	U	U	U	U	U	U	U	U	U
2	15	14	U	U	U	U	U	U	U	U	U	U	U	U
3	29	10	LS	LS	U	LS	LS	U	LS	U	U	U	LS	U
4	39	9	LS	LS	U	LS	LS	U	LS	U	U	U	LS	U
5	51	6	LS	LS	U	LS	LS	U	LS	U	U	U	LS	U
6	57	9	U	U	LS	U	LS	U	LS	LS	U	U	U	U
7	66	6	LS	U	LS	U	LS	U	LS	LS	U	U	LS	LS
8	72	10	LS	U	LS	U	LS	U	LS	LS	U	U	LS	LS
9	82	11	LS	U	U	LS	U	LS	LS	U	U	U	LS	LS
10	93	7	U	U	U	LS	U	LS	LS	U	LS	LS	U	U
11	100	11	LS	U	U	LS	U	LS	LS	U	LS	LS	U	LS
12	111	7	LS	U	LS	U	U	LS	LS	U	LS	LS	U	LS
13	118	11	LS	U	LS	U	U	LS	LS	U	LS	U	U	LS
14	129	2	U	U	U	U	U	U	U	U	U	U	U	U

Note: DAS = days after sowing; Days = days in interval; U = uniform irrigation of 50 mm.; and LS = linesource irrigation.

V - ESTIMATION OF AN INDEX OF WATER STRESS

As stated at the beginning of the previous section, an important part of identifying potential risk reducing genotypes for specific locations is generating a set of observations on yields by location. In this study, observations on yield for each genotype in each of three selected locations are generated in two steps by combining the experimental response to water stress with location-specific information on environmental factors. The first step relates the patterns of drought observed in each sample location to those in the trial, taking account of all three parameters of drought: timing, duration and intensity. In the second step, the relationship between yield and different patterns of stress is estimated; location-specific yield distributions can then be simulated. In this section, an independent measure of water stress is derived, and the distribution of drought conditions in each location is described. The specification of the response relationship between genotype yields and water stress and the simulation of yield distributions are discussed in section *VI*.

Development of a Measure of Water Availability

To control the drought treatments through the application of irrigation without interference from rainfall, the trial was planted in the late post rainy season and continued into the summer season under conditions that, meteorologically, are different from those in the kharif (rainy season). Furthermore, the trial is characterized by discrete applications of water, irrigation being applied only when symptoms of wilting were observed in the non-stressed control plots. This is in contrast to the essentially continuous nature of rainfall in the rainy season, when soil moisture may be continually replenished by rain showers. Thus, the problem of simulating yields in each of the three sample locations is compounded by the difficulties involved in translating results from the post rainy season to rainy season conditions.

To do this, two strategies were tried. The first involves matching each weather pattern to a specific pattern in the trial, using the pattern specific measure of cumulative percent water deficit (%WD) discussed by Nageswara Rao and Williams (1985), Nageswara Rao *et al.* (1985) and Williams *et al.* (1986). The results from both strategies are reported in detail by Bailey (1988) but those from the first strategy were extremely disappointing. There was little correlation between the patterns in the experiment and the percent water deficits calculated from the weather data for each site. Furthermore, while the measure of %WD takes account of the differences in evaporative demand between the two seasons, it does not take account of the possible carryover of soil moisture from one interval to the next. This is inherent in the design of the experiment, irrigations being timed according to when soil moisture has been depleted to the point where non-stressed plants show signs of wilting.

In the second strategy, an index of relative water availability (RAW, available moisture relative to potential water requirements) is developed based on a simple daily soil moisture budgeting approach. It takes account of the soil moisture holding capacity and the consumptive water use requirements of the plant, given potential evapotranspiration rates and stage of crop growth, and also allows for the carryover of soil moisture from one period to another.

Evapotranspiration (ET) is the sum of the water evaporated from a moist soil surface and foliage, without passing through the plant, and the soil water used by the plant in transpiration. Crop dry matter production is closely linked to the amount of water transpired by the crop (Campbell and Diaz 1986). Potential evapotranspiration (PET) is an indicator of the possible maximum consumptive use when water is not limiting and there is full foliage ground cover (Baier and Robertson 1966). It has been maintained that evaporation from a freely evaporating open water surface provides an upper limit on PET (Arnon 1975), and that "ET from a field fully covered with lush green foliage does not exceed evaporation from a large open body of water in its neighborhood" (Minhas *et al.* 1973, p.384). Potential evapotranspiration can therefore be viewed from two perspectives. On the one hand, it is the amount of water potentially required to meet ET demands. On the other, it can be taken to represent the totality of the prevailing weather conditions, "it is an index composed of the various determinants of evaporation" (Minhas *et al.* 1973, p.384).

When water is freely available, plant development is determined by chronological time (McGuckin *et al.* 1987), and evapotranspiration will proceed at the potential rate, but when water is limited, plant development will be determined by other factors. There are complex interactions among the plant, soil and meteorological factors which determine how a plant reacts to a given moisture regime. As a soil moisture deficit accumulates, transpiration is progressively restricted, reducing ET. Under conditions of high temperature, low relative humidity, high winds, etc., both PET and the plant's demand for water increase and the plant will be more sensitive to moisture deficits. Estimating the amount of water available for ET, relative to PET, gives a measure of the shortfall in water required by the plant.

The amount of water available for transpiration is determined by the amount of water left in the soil after other demands for water have been satisfied. Water is supplied by precipitation and irrigation. If water is supplied at a rate faster than it can infiltrate the soil surface, then it may be lost to run-off. Water that enters the soil may drain beyond the root zone. The amount lost to such deep drainage depends on the soil structure and its moisture storing capacity. Only that water which is stored in the root zone is available for transpiration.

The problem is that of estimating the soil moisture balance remaining at the end of each interval that, along with any water added by rainfall or irrigation less any water lost to run-off or to deep percolation beyond the root zone, will be available to the crop in the following interval. The soil moisture balance and the rate of ET are inextricably linked. On any given day, consumptive water use not only depends on the soil water available on that day, but also determines the soil water available for use the next day. Daily consumptive water use also depends on the stage of crop growth which is a function not only of time, but also of the preceding soil water conditions and the degree to which ET demands have been satisfied.

There is a vast literature in the agricultural sciences dealing with the determinants of the soil-water-plant system. Models of various degrees of sophistication have been developed; they range from simple soil water bookkeeping methods to simulation models which incorporate the physiological processes of the plant with detailed information on meteorological and soil factors. Much of this literature was reviewed in developing a measure of relative water availability (see Bailey 1988). A number of simplifying assumptions have been made in the interests of minimizing meteorological and soil data requirements.

A modification of the simple soil moisture budgeting approach is adopted. The initial soil moisture balance is determined, inputs of water in the form of precipitation or irrigation, less any water lost to run-off or deep drainage, are added, the estimated evapotranspiration is subtracted and the balance carried forward into the next period. The soil moisture at the end of a given interval of time, *i*, is then expressed as:

$$SW_i = SW_{i-1} + P_i + I_i - Ro_i - Dr_i - ET_i \quad (20)$$

where SW_i = the soil moisture at the end of interval i, P_i = precipitation in interval i, I_i = irrigation in interval i, Ro_i = run-off in interval i, Dr_i = deep drainage in interval i, and ET_i = evapotranspiration in interval i.

In estimating ET, many approaches rely on the use of experimentally determined relationships between ET and PET. A common approach is the use of a crop factor or coefficient, k, such that ET = kPET, where k varies with stage of growth. Stern (1986) applied this approach to rainfall data for Hyderabad to derive water requirements of groundnuts, in 10-day intervals, over the kharif growing season; his crop coefficient increases from 0.3 at planting to 1.0 at 60-90 DAS, declining to 0.6 at maturity. Such crop coefficients are, however, specific to the experimental conditions. More important, ET will vary with available soil moisture and, therefore, k has to be corrected when soil moisture is limited (Hanks and Rasmussen 1982). Furthermore, the crop coefficient is specific to the stage of crop growth which is partially determined by past soil moisture conditions and ET rates.

The use of crop coefficients highlights the difference in objectives in many of the studies reviewed. If the objective is to determine optimal irrigation scheduling, then the crop factor approach is used to determine the amount of water required by the crop at any stage given PET demands. Our objective, on the other hand, is to approximate the actual soil water status in any given interval of the growing season. Consequently, the evapotranspiration estimated by kPET is interpreted as the potential water requirements (PWR = kPET) of the crop, *i.e.*, it is the amount of water required by the crop for evapotranspiration under optimal soil water conditions and given the stage of growth of the crop. This can be viewed as the amount of water potentially required by the crop when water is freely available, and has been freely available over preceding intervals so that the crop has developed at the optimum rate. Following Stern (1986), it is assumed that evapotranspiration will proceed at the maximum rate possible, given the available soil moisture.

The daily soil water budget is estimated as follows:

1. The available soil moisture capacity (SWC) is 100mm.

2. Daily PET is approximated by daily pan evaporation data.

3. Potential water requirements during any day, t,

(21) $PWR_t = k_t\ PET_t$,

where k_t is taken from Stern (1986).

4. The water available for evapotranspiration on day t (AW_t) is the sum of the soil moisture balance from the previous day (SW_{t-1}) and the water added by precipitation (P_t) or irrigation (I_t), less any losses to run-off or deep drainage (L_t):

(22) $AW_t = SW_{t-1} + P_t + I_t - L_t$.

It is assumed that losses (L_t) occur whenever the water added, plus that already in the soil, exceeds the storage capacity (SWC), *i.e.*,

(23) if $SW_{t-1} + P_t + I_t > SWC$ then $L_t = (SW_{t-1} + P_t + I_t) - SWC$ and $AW_t = SWC$,

if $SW_{t-1} + P_t + I_t \leq SWC$ then $L_t = 0$ and $AW_t = SW_{t-1} + P_t + I_t$.

It is not possible to differentiate between losses to run-off and to deep percolation. A more sophisticated approach would take account of the rate of rainfall and infiltration rates.

5. The soil moisture balance at the end of day t (SW_t) depends on the estimated rate of evapotranspiration (ET_t) which is determined by the water available (AW_t) and the potential water requirements (PWR_t). If available moisture exceeds PWR, then it is assumed that $ET_t = PWR_t$ and the surplus is stored as soil moisture; if PWR exceeds available moisture, then it is assumed that soil moisture is depleted to zero as follows:

(24) if $AW_t > PWR_t$ then $ET_t = PWR_t$ and $SW_t = AW_t - PWR_t$

otherwise $ET_t = AW_t$ and $SW_t = 0$.

The soil water balance, SW_t, is then carried forward and used in the estimation of the following day's water budget.

To begin the soil moisture budgeting procedure, an initial soil moisture balance must be determined. In the absence of plant cover, evaporation from a bare soil surface is equal to 30%-40% of pan evaporation (Boote *et al.* 1982). This fact is employed in the determination of initial soil moisture levels at date of planting (SW_0). In the trial, planted in the late post rainy season, initial soil moisture, before the application of the first uniform irrigation, is assumed to be zero. In the kharif, soil moisture is assumed to be zero until the first rains occur; daily ET is then fixed at 40% of pan evaporation and the same rules as above applied to determine the soil moisture balance on the day of planting.

Daily mean air temperatures during the post rainy - summer season are very different to those in the kharif. Plant development is predominantly controlled by temperature; there are conspicuous differences in the time to flowering, podding and the total duration of growth of groundnuts between the two seasons (Ong 1986). The 130-day growing period in the trial is appropriate for the late post rainy - summer season. In the kharif, however, the normal time to maturity of the genotypes included in the trial would be about 105 days. The time to maturity of the experimental crop is some 25 days longer in duration due mainly to the lower temperatures during December, which delay germination, and to the shorter day length, which delays plant development.

A 105-day growing season was defined for each year for which data were available in each of the sample sites. Planting dates in each year were determined based on farming practices and the actual rainfall received. In two years in Hyderabad and 15 years in Anantapur, it was determined that there was insufficient rainfall to allow planting of groundnuts. In such years, farmers would have planted an alternative crop, such as pigeonpea. The 105-day growing seasons were then divided into intervals corresponding to the irrigation intervals in the trial. The first 29 days in the trial, which were irrigated uniformly, "telescope" into about 15 days in the kharif due to the higher temperatures in the early kharif promoting earlier germination. The remaining 90 days of the 105-day growing season are divided into 11 intervals of eight days each, corresponding to the 11 irrigation intervals in the trial, plus two remaining days which correspond to the 14th, pre-harvest, interval in the trial.

To account specifically for the effect of timing of water application, the growing seasons were further divided into four growth phases: G1: seedling-flowering vegetative phase: intervals 3 - 5, G2: pegging to beginning of podset: intervals 6 - 8, G3: podset to podfilling: intervals 9 - 11, and G4: podfilling to maturity: intervals 12 and 13. The differences in timing and length of the growth stages in the two seasons correspond to the differences in time to flowering and podding, and in the total duration of growth, noted by Ong (1986).

A measure of the total available soil moisture during each of these growth phases (AW_i; $i=1,2,3,4$) is obtained from the daily water budgets. By assuming that evapotranspiration proceeds at the potential rate, unless soil water is limiting in which case the amount of water used in evapotranspiration is assumed to equal the amount available, AW_i is the sum of the estimated daily evapotranspiration for the days in that growth stage:[4]

(25) $$AW_i = \sum_{t \in i} ET_t \,.$$

Relative available water (RAW_i) is then defined as available water relative to potential water requirements:

(26) $$RAW_i = AW_i/PWR_i, \text{ where } PWR_i = \sum_{t \in i} PWR_t \,.$$

RAW_i is effectively constrained to lie between zero and unity as follows:

if $AW_t \geq PWR_t$ for all $t \in i$, then $\sum_{t \in i} ET_t = \sum_{t \in i} PWR_t$ and $RAW_i = 1$;

if $AW_t < PWR_t$ for any $t \in i$, then $ET_t < PWR_t$, $\sum_{t \in i} ET_t < \sum_{t \in i} PWR_t$ and $RAW_i < 1$;

if $AW_t = 0$ for all $t \in i$, then $\sum_{t \in i} ET_t = 0$ and $RAW_i = 0$.

RAW_i is a measure of the degree to which the crop's requirements for water during each growth phase are satisfied. It is not proposed that this approach gives an accurate representation of the soil water regimes in the different sites and seasons; it simply provides a means of determining a relative measure of available water that is comparable across sites and seasons. A value of $RAW_i = 1$ indicates that sufficient water was available to meet the potential water requirements of the crop during the interval, a value less than unity that available water was insufficient to meet potential water requirements. No notion of the degree of stress experienced by the plant is attached to the values of RAW_i. RAW_i is simply an index of relative water availability; whether or not the crop experienced stress is revealed by relating the yields achieved in each pattern to the values of RAW_i in each interval.

Distributions of RAW

Table 6 presents summary statistics on relative water availability in each of the four growth phases in the experiment and each location. Due to the low temperatures during the early part of the experiment, PET, and therefore PWR, during the early growth phase were relatively low. Because the experiment was adequately irrigated during the first 28 days

---

[4] AW_i can also be defined as the sum of the soil water balance on the last day of the previous growth stage and the total precipitation or irrigation, less the amount lost during the interval and the amount remaining in the soil at the end of the interval which was in excess of requirements and is available in the subsequent interval:

$AW_i = SW_{i-1} + \sum_{t \in i} P_t + \sum_{t \in i} I_t - \sum_{t \in i} L_t - SW_i$. Expanding, it can be shown that

$AW_i = \sum_{t \in i} AW_t - SW_i = \sum_{t \in i} ET_t \,.$

Table 6. Summary Statistics on Distributions of RAW

RAW	Growth Phases G1	G2	G3	G4
Experiment				
Maximum	1.000	1.000	0.716	0.790
Minimum	0.497	0.012	0.024	0.005
Mean	0.935	0.795	0.507	0.570
Std. dev.	0.148	0.282	0.189	0.269
Skewness	-2.105	-1.289	-0.969	-0.936
P[RAW ≤ 0.5]	0.01	0.18	0.39	0.33
Hyderabad				
Maximum	1.000	1.000	1.000	1.000
Minimum	0.394	0.278	0.367	0.000
Mean	0.977	0.934	0.893	0.923
Std. dev.	0.092	0.139	0.166	0.200
Skewness	-5.172	-3.062	-1.521	-3.220
P[RAW ≤ 0.5]	0.01	0.02	0.04	0.05
Anantapur				
Maximum	1.000	1.000	1.000	1.000
Minimum	0.000	0.000	0.000	0.000
Mean	0.441	0.352	0.453	0.564
Std. dev.	0.299	0.287	0.317	0.392
Skewness	0.468	0.679	0.523	-0.274
P[RAW ≤ 0.5]	0.60	0.73	0.64	0.41
Gujarat				
Maximum	1.000	1.000	1.000	1.000
Minimum	0.554	0.033	0.008	0.000
Mean	0.986	0.944	0.793	0.603
Std. dev.	0.079	0.188	0.302	0.470
Skewness	-5.654	-4.168	-1.200	-0.392
P[RAW ≤ 0.5]	0	0.03	0.25	0.41

before the drought patterns were imposed, stored soil moisture was sufficient to maintain relatively high levels of AW during the early growth phase even in plots that received little or no water, and consequently $0.5 \leq RAW_1 \leq 1$. RAW_2 covers the full range from virtually zero, in the highest intensities of patterns 5 and 7 (*P5:8* and *P7:8*), to unity in adequately irrigated treatments. As PET increased over the experimental season and the crop matured (with corresponding higher crop coefficients, k_i), PWR increased and this is reflected in RAW in growth stages G3 and G4: $RAW_3 \leq 0.72$ and $RAW_4 \leq 0.79$. The result is that the distribution of RAW, even in the later growth phases, is negatively skewed.

In each year in the three locations the total rainfall, the total water lost to run-off or deep percolation, the total water available to the crop, the total potential water requirements, and the total number of drought days in Hyderabad and Anantapur, or drought intervals in Gujarat, were accumulated over the eleven intervals, where drought days (intervals) are defined first, as any day (interval) in which soil moisture is zero, and second, as any day (interval) in which soil moisture is 25% of soil capacity or less. Data are reported in Bailey (1988), but the important differences across sites are summarized below.

In Hyderabad, the total amount of rain received during the eleven intervals was highly variable, ranging from 154 mm. to 990 mm., with a mean value of 446 mm. Similarly, total losses were highly variable; high losses were, however, associated with high rainfall and, consequently, there was relatively little variation in total available water which in many years was sufficient to meet total water requirements. This fact is reflected in the mean values and negatively skewed distribution of RAW in the four growth phases in Table 6.

Total water requirements in Anantapur were far higher than in Hyderabad due to higher temperatures, and therefore higher PET rates, while total rainfall was lower, ranging from only 53 mm. to 523 mm., with a mean of 286 mm. Levels of available water reflected the variability in rainfall and, on average, provided only 42% of the total water required. As can be seen from Table 6, the distributions of RAW in the four growth phases in Anantapur are very different to those in Hyderabad. Mean values are far lower, the probability of RAW falling below 0.5 far higher and, except for growth phase G4, the distributions are positively skewed.

In Gujarat, the variation in total rainfall was associated with date of planting. Years in which planting occured later in the season showed lower total rainfall levels. While mean rainfall was high, total losses were also high; high water losses, however, generally occured in years of high rainfall so that even with such high losses, total available water was often sufficient to meet total water requirements. Years in which a number of intervals were droughted tend to be either those with low rainfall or those in which planting was delayed by the late onset of rain; the exception is 1979 which received 1,673 mm. rain but lost most of it to run-off or deep drainage. Mean values of RAW in each growth phase are similar to those in Hyderabad, except for G4 (Table 6). The shorter rainy season in Gujarat increases the probability of late season droughts, consequently the distribution of RAW_4 is less negatively skewed than distributions of RAW_1 - RAW_3.

VI - SIMULATION OF YIELD DISTRIBUTIONS

Having derived an index of relative water availability, RAW, the relationship between yields in the trial and RAW must be estimated to allow simulation of yields in each year in each study site.

Yield Response

In examining the experimental data, it was discovered that yields recorded for intensity 8 of pattern 8 duplicated those recorded for intensity 7 of pattern 8. This was attributed to errors in recording or data entry and therefore one set of observations was deleted from the data set. Due to these deletions, plus a number of other missing observations and suspected recording errors, yields in the experiment were averaged over replicates. Mean pod yields (over all genotypes) for each intensity of each pattern are presented in Table 7.

Our objective is to estimate the relationship between yields and RAW through the specification of a response function for each genotype such that:

(27) $Y_{gpd} = f_g \, (RAW_{pdi})$

Table 7. Mean Yields (gm./m.^{2}) by Pattern and Drought Intensity

Pattern[a]	Drought Intensity 1	2	3	4	5	6	7	8
P1	319.0 (45.9)[b]	272.7 (46.7)	263.1 (35.1)	230.7 (37.7)	175.8 (32.6)	130.2 (28.3)	91.7 (22.5)	38.1 (11.8)
P2	328.1 (48.2)	306.1 (48.1)	316.1 (61.3)	280.3 (65.6)	305.0 (51.8)	280.8 (48.5)	290.7 (52.2)	255.9 (44.3)
P3	286.4 (52.5)	287.8 (63.2)	212.4 (36.3)	191.7 (31.2)	173.4 (24.7)	123.8 (27.0)	77.0 (15.2)	87.6 (24.0)
P4	219.3 (38.5)	263.1 (41.5)	266.0 (46.2)	227.0 (45.0)	182.8 (28.0)	141.9 (26.8)	108.8 (19.2)	77.3 (19.3)
P5	272.7 (50.9)	278.4 (64.7)	291.0 (46.3)	267.2 (49.1)	251.8 (44.9)	196.3 (39.8)	153.4 (33.9)	153.8 (28.1)
P6	256.3 (37.0)	244.4 (35.6)	205.2 (35.6)	155.9 (31.9)	113.3 (19.4)	89.5 (22.0)	56.7 (12.9)	37.4 (11.8)
P7	269.4 (38.9)	269.3 (33.6)	200.3 (31.2)	131.3 (24.8)	48.8 (19.9)	19.4 (14.0)	2.5 (4.7)	6.9 (11.9)
P8	301.8 (51.9)	265.7 (62.3)	267.9 (45.1)	229.9 (36.7)	205.8 (38.0)	172.5 (39.6)	128.4 (37.5)	. .
P9	263.4 (51.0)	250.7 (47.9)	210.8 (42.6)	217.0 (33.8)	195.7 (31.3)	142.4 (30.7)	106.4 (22.2)	82.3 (17.8)
P10	266.9 (52.0)	249.5 (42.5)	231.2 (37.2)	214.4 (35.9)	161.9 (41.4)	102.2 (32.9)	84.3 (37.3)	104.4 (36.0)
P11	293.3 (45.6)	288.0 (43.1)	318.3 (58.1)	290.3 (52.7)	219.3 (36.8)	166.4 (29.4)	156.8 (26.3)	101.1 (22.0)
P12	267.3 (52.3)	244.3 (51.3)	210.2 (28.6)	179.7 (23.3)	159.2 (30.7)	102.9 (21.7)	67.0 (15.9)	32.3 (15.6)

[a]See sections IV and V for a discussion of the drought patterns and intensities.

[b]Figures in parentheses are standard deviations.

where Y_{gpd} = the average yield (over replicates) of the g^{th} genotype in the d^{th} drought intensity of the p^{th} pattern, and RAW_{pdi} = the average relative water availability in the i^{th} growth phase in the d^{th} drought intensity of the p^{th} pattern of the experiment, $p = 1, \ldots, 12$, $d = 1, \ldots, 8$, *and* $i = 1, \ldots, 4$. In each location, yields in any year, y, can then be simulated by using the estimated function and the calculated values of RAW at the sites:

(28) $\hat{Y}_{gy} = \hat{f}_g\,(RAW_{yi})$

where $\hat{Y}_{gy}$ = estimated yield of the g^{th} genotype in the y^{th} year, $\hat{f}_g$ = the estimated response function, RAW_{yi} = relative water availability in the i^{th} growth phase in year y.

In their initial analysis of the experimental results, ICRISAT scientists used a pattern specific measure of water deficit (%WD) to compare the effects on average yields of different patterns of drought. The index of relative available water, $RAW = AW/PWR$, is analogous to 1 - %WD. Thus, the general results on average crop response to %WD reported by Nageswara Rao and Williams (1985), Nageswara Rao *et al.* (1985) and Williams *et al.* (1986) are useful in developing some initial hypotheses regarding the response to RAW. The simplest hypothesis is that yield is a linear function of relative water availability in each interval. Such an assumption, however, implies that yield increases monotonically with RAW and never reaches a maximum (a biological impossibility), imposes a constant marginal productivity of RAW in each interval, and does not allow for possible interactions between the degree of relative water availability in different intervals.

In contrast, Nageswara Rao *et al.* (1985) found that interactions between different stages of drought did exist; depending on whether or not the early vegetative phase was droughted, response to subsequent droughts was modified. They also observed the effect on yields of water stress in the early growth phase alone. Average yields do not decline linearly with increasing stress in the early growth stage. This may be attributed to the groundnut plants' capacity, at this stage of growth, to lie dormant until soil water is replenished. From an examination of the disaggregated results from the trial, there are indications that, over some range, yields may in fact increase with increasing water deficits in the early growth stage, *i.e.*, over some range, the marginal productivity of RAW_1 may be negative. For all other growth phases, it is hypothesised that as RAW_i approaches zero and the plant is increasingly stressed, a small increase in AW_i relative to PWR_i will result in a relatively large marginal response, *i.e.*, as $RAW_i \rightarrow 0$, f_i is large. However, as RAW_i approaches unity, the plant has adequate water and responses to a small increase in AW_i will be negligible, *i.e.*, as $RAW_i \rightarrow 1$ then $f_i \rightarrow 0$.

Assessment of Response Relationship

Modelling the relationship between RAW and yield proved to be the most problematic part of the analysis. Not only is one concerned with the ability of the model to represent the experimental results accurately, but one must also deal with the problem of predicting outside the range of RAW represented in the trial. As discussed in section *V*, the full range of $0 \leq \text{RAW} \leq 1$ is not covered in all growth phases in the trial.

Based on these considerations, a number of model specifications were compared on the basis of a) their ability to model accurately the yield response within the range of RAW in the trial, and b) their performance in predicting yields in years with drought patterns not included in this range.

Assessment of goodness of fit of the alternative specifications of the functional relationship between yield and RAW_1 - RAW_4 within the experimental range was based on a number of criteria. Actual and predicted yields were compared and the size and pattern of residuals examined. Recalling, from section *III*, that results from any stochastic dominance analysis are extremely sensitive to any errors in the measurement of the lower tails of the CDFs, particular attention was paid to the model's predictive capabilities in the lower yielding environments in the trial.

Second, predicted yields were regressed on actual yields; an intercept near zero, a slope coefficient approaching unity and a low MSE indicate a reasonable correspondence between actual and predicted yields. The predictive performance of a model can also be assessed using Theil's inequality coefficient:

$$(29) \quad U = \frac{\sqrt{1/T \sum_{t=1}^{T} (Y_t^P - Y_t^A)^2}}{\sqrt{1/T \sum_{t=1}^{T} (Y_t^P)^2} + \sqrt{1/T \sum_{t=1}^{T} (Y_t^A)^2}}$$

where Y_t^A = actual yield, Y_t^P = predicted yield (Pindyck and Rubinfeld, 1981, p.364). If $U = 0$, then $Y_t^P = Y_t^A$ for all t and there is a perfect fit. Pindyck and Rubinfeld (1981, p.365) show that the Theil coefficient can be decomposed in the following manner:

$$(30) \quad 1/T \sum_{t=1}^{T} (Y_t^P - Y_t^A)^2 = (\overline{Y}_t^P - \overline{Y}_t^A)^2 + (\sigma^P - \sigma^A)^2 + 2(1 - \rho)\sigma^P\sigma^A$$

where $\overline{Y}^P$, $\overline{Y}^A$, σ^P, and σ^A are means and standard deviations of the predicted and actual yields respectively. They then define three proportions of inequality: bias, variance, and covariance. The bias proportion, U^M,

$$(31) \quad U^M = \frac{(\overline{Y}^P - \overline{Y}^A)^2}{(1/T) \sum_t (Y_t^P - Y_t^A)^2}$$

is an indication of systematic error, measuring the extent to which the average values of the actual and predicted yields deviate from each other; U^M values close to zero are desirable. The variance proportion U^S,

$$(32) \quad U^S = \frac{(\sigma^P - \sigma^A)^2}{(1/T) \sum_t (Y_t^P - Y_t^A)^2}$$

indicates the ability of the model to replicate the degree of variability in yields in the experiment; again low values are desirable. The covariance proportion, U^C,

$$(33) \quad U^C = \frac{2(1 - \rho)\sigma^P\sigma^A}{(1/T) \sum_t (Y_t^P - Y_t^A)^2}$$

measures the unsystematic error and is calculated as the remainder: $U^C = 1 - U^M - U^S$; for any value of $U > 0$, the ideal distribution of inequality over the three sources is $U^M = U^S = 0$, and $U^C = 1$.

Assessment of a model's performance in predicting yields outside the experimental range was initially based on the computation of predicted yields for three hypothetical cases:

a) $RAW_i = 1$ for all i = 1,...,4;

b) $RAW_1 = 0.1$, $RAW_j = 1$, $j \neq 1$;

c) $RAW_1 = 0.1$, $RAW_2 = 1$, $RAW_3 = 0.72$, $RAW_4 = 0.79$.

In all three cases yields are expected to be relatively close to the maximum in the experiment. If yield response to these hypothetical drought patterns met *a priori* expectations, then yields were predicted for all years in each location and examined closely for inconsistencies.

Specification of Response Functions

In finding the appropriate response relation, a number of model specifications were examined. These included: a) a general power function; b) a generalized quadratic and a third order polynomial; c) a function proposed by Minhas *et al.* (1973) where marginal products go from ∞ to zero as the input goes from zero to unity; d) a function proposed by Mitscherlich (in Hexem and Heady, 1978) which states that yields can be expanded through increasing levels of any single growth factor as long as that factor is not present in sufficient amounts to produce maximum yields; e) a generalized logistic growth curve and f) a Gompertz growth curve.

The results from the experimentation with these functions are reported in detail in Bailey (1988) In general, however, these functions were too inflexible to model the lower tails of the distribution and did not predict well outside the range of the data, or the nonlinear estimation process did not converge.

As an alternative, a flexible form, the transcendental logarithmic (translog) production function was estimated. Letting $Y_{gpd} = Y$ and $RAW_{pdi} = X_i$ from equation (27), then :

(34) $$Y = a_0 \prod_i X_i^{a_i} \prod_i X_i^{1/2[\sum_j b_{ij} lnX_j]}$$

or in logarithmic form:

(35) $$lnY = lna_0 + \sum_i a_i lnX_i + 1/2 \sum_i \sum_j b_{ij} lnX_i lnX_j .$$

(36) The production elasticity $e_i = dlnY/dlnX_i = a_i + \sum_j b_{ij} lnX_j$,

(37) $$f_i = dlnY/dlnX_i.Y/X_i = (a_i + \sum_j b_{ij} lnX_j).Y/X_i = e_i.Y/X_i,$$

(38) $f_{ii} = [b_{ii} + (e_i - 1)e_i].Y/X_i^2$ and

(39) $f_{ij} = [b_{ij} + e_i e_j].Y/X_i X_j$.

The elasticities of substitution between inputs, e_{ij}, are not constant, but vary with the level of inputs (Boisvert, 1982). Recalling that $0 \leq X_i \leq 1$, then $lnX_i \leq 0$. The marginal productivity, f_i, can be positive for a range of values of X_j if $b_{ij} < 0$, but can also be negative if $b_{ij} > 0$. Positive marginal productivity requires that $e_i > 0$. If $e_i < 1$ then $f_{ii} < 0$ (and marginal productivity is diminishing) if $b_{ii} < 0$ or if $b_{ii} > 0$ and $b_{ii} < |e_i(e_i - 1)|$; if $e_i > 1$, then $f_{ii} < 0$ only if $b_{ii} < 0$ and $|b_{ii}| > (e_i - 1)e_i$. When all $X_i = 1$ then $lnX_i = 0$, $e_i = a_i$ and, as $Y = a_0$, $f_i = a_i a_0$. The sign of f_i will be determined by the sign on a_i.

A translog production function was estimated for each genotype in the trial by single equation OLS. Within the experimental range, predicted yields were close to actual yields for all patterns, particularly in the lower yielding patterns, and the function did not overestimate yields in the higher intensities of pattern P2 (early drought), a problem encountered with many of the other functions that were tried. Regressions of predicted yields on actual yields produced intercepts close to zero, slope coefficients less than, but reasonably close to, unity and $\overline{R}^2$ in the range 0.8 - 0.9, except for genotype GNO16.

Multicollinearity was an evident problem, as exhibited by large standard errors on estimated coefficients particularly for terms involving X_1 and, because X_1 rarely fell far below unity, for crossproducts involving X_1. One potential solution is to remove selectively those squared and crossproduct terms whose t-ratios are below some critical value. However, without any *a priori* rationale for deleting terms, this strategy could ultimately destroy the flexibility of the function (Boisvert 1982). One of our concerns is with the prediction of yields when X_i's are outside the range in the experiment, and particularly when $X_1 \rightarrow 0$. In the initial estimation of the full translog models, there were differences between genotypes in the sign and magnitude of estimated coefficients on terms involving X_1. As $X_1 \rightarrow 0$, $lnX_1 \rightarrow -\infty$ and $(lnX_1)^2 \rightarrow \infty$. A large positive coefficient on $(lnX_1)^2$, relative to that on lnX_1, or a large negative coefficient on lnX_1, relative to that on $(lnX_1)^2$, result in high predicted yields. Therefore, for those genotypes that had large, positive, insignificant coefficients on $(lnX_1)^2$, the term was deleted. GNO6 had a large, negative insignificant coefficient on lnX_1, relative to $(lnX_1)^2$, and this term was deleted.

Coefficients, and their associated standard errors, for the log formulation of the final translog models are given in Table 8. The computed Theil coefficients for predicted yields in the trial for all 22 genotypes are given in Table 9.

The translog is attractive because of its flexibility; it places no restrictions on marginal productivity and the elasticities of substitution between inputs are not constant. Because of this flexibility it is difficult to make general statements regarding response to RAW in different growth phases. Marginal products were estimated from the final models for various values of RAW_1 - RAW_4. These are presented in Appendix B.

When RAW_1 - RAW_4 are set at their mean values the marginal response to RAW in growth phases 2, 3 and 4 is positive for all genotypes $(f_i > 0,\ i = 2, 3, 4)$ while in growth phase 1 it is negative $(f_1 < 0)$ for all genotypes except GNO22; this supports our earlier hypothesis that yields may be increased by some degree of stress in the early phase.

When RAW_1 - RAW_4 are set at their maximum values, $f_i > 0$ for $i = 2, 3, 4$. In this case, we are particularly concerned with the marginal products of RAW_3 and RAW_4 which have implications for the simulation of yields in years when values of RAW in growth phases 3 and 4 exceed the maximum in the trial; while yield response to RAW_3 and RAW_4, at their limits in the trial, is increasing, it is increasing at a decreasing rate $(f_{33}, f_{44} < 0)$. At maximum values of RAW, the marginal response to RAW_1 is generally negative with the exception of GNO20 and GNO22, which have large positive marginal products and, to a lesser extent, GNO2 and GNO15.

Two other scenarios are specified in Appendix B. First, by letting RAW_1 = 0.5, its minimum value in the trial, and RAW_2 - RAW_4 = their maximum values, we can compare the results with those from the previous scenario and gauge the effect of water stress in the early phase only. Genotypes differ in their response to a change in RAW in the early phase.

Table 8. Estimated Coefficients for Translog Models, Estimated by Ordinary Least Squares

Independent Variables, log form[b]	Dependent Variable, Yg[a]					
	Y1	Y2	Y3	Y4	Y5	Y6
Intercept	6.1026	6.2786	6.3139	6.4534	6.0712	6.3450
	(0.203)	(0.213)	(0.132)	(0.140)	(0.105)	(0.129)
$\ln X_1$	0.2876	0.6202	0.1772	0.1891	-1.4929	.
	(1.036)	(1.087)	(1.161)	(0.716)	(0.922)	.
$\ln X_2$	0.3424	0.3665	0.2879	0.3551	0.5990	0.5788
	(0.286)	(0.300)	(0.193)	(0.197)	(0.153)	(0.171)
$\ln X_3$	0.6480	0.8257	1.1426	1.1258	1.0598	1.1106
	(0.587)	(0.615)	(0.380)	(0.405)	(0.302)	(0.361)
$\ln X_4$	0.7976	0.5629	0.4180	0.4350	0.1861	0.4161
	(0.313)	(0.328)	(0.201)	(0.216)	(0.159)	(0.198)
$(\ln X_1)^2$	.	.	-0.4711	.	-2.7672	0.2811
	.	.	(0.761)	.	(1.398)	(1.014)
$(\ln X_2)^2$	0.1459	0.1196	0.0445	0.1658	0.0906	0.0808
	(0.196)	(0.205)	(0.126)	(0.135)	(0.100)	(0.1260)
$(\ln X_3)^2$	-0.1878	-0.0884	0.1323	0.0734	0.2570	0.1470
	(0.386)	(0.405)	(0.248)	(0.267)	(0.197)	(0.235)
$(\ln X_4)^2$	0.0083	0.0172	0.0045	0.0057	0.0387	0.0355
	(0.067)	(0.071)	(0.043)	(0.047)	(0.034)	(0.043)
$\ln X_1.\ln X_2$	-0.0465	0.2233	0.3903	-0.5579	0.2268	0.2394
	(1.169)	(1.2260	(0.775)	(0.807)	(0.616)	(0.779)
$\ln X_1.\ln X_3$	1.8515	1.4640	0.6396	0.6876	-0.3868	0.1923
	(1.735)	(1.820)	(1.111)	(1.198)	(0.882)	(1.045)
$\ln X_1.\ln X_4$	-0.5317	-0.0226	-0.0384	0.2344	0.4555	-0.0108
	(0.674)	(0.707)	(0.432)	(0.466)	(0.343)	(0.420)
$\ln X_2.\ln X_3$	-1.4236	-1.2058	-0.6504	-0.3783	0.2074	-0.1942
	(0.396)	(0.415)	(0.254)	(0.273)	(0.202)	(0.252)
$\ln X_2.\ln X_4$	0.4839	0.2566	0.1905	-0.0142	-0.1146	0.0851
	(0.328)	(0.344)	(0.210)	(0.227)	(0.167)	(0.211)
$\ln X_3.\ln X_4$	0.3655	0.2026	0.1048	0.1600	-0.2080	0.0020
	(0.290)	(0.304)	(0.186)	(0.200)	(0.147)	(0.181)
$\overline{R}^2$	0.8646	0.8582	0.8685	0.8292	0.8227	0.810

Table 8. (continued)

Independent Variables, log form[b]	Dependent Variable, Yg[a]					
	Y7	Y8	Y9	Y10	Y11	Y12
Intercept	5.6116 (0.100)	6.5250 (0.272)	5.8690 (0.077)	5.7612 (0.102)	6.3598 (0.137)	5.9725 (0.103)
$\ln X_1$	-1.4953 (0.880)	0.0800 (0.393)	-0.3845 (0.672)	-2.1701 (0.892)	-0.4353 (1.200)	-0.1359 (0.901)
$\ln X_2$	0.1488 (0.146)	0.6243 (0.384)	0.3764 (0.112)	0.3438 (0.148)	0.5708 (0.200)	0.6096 (0.150)
$\ln X_3$	0.0586 (0.288)	1.7864 (0.789)	0.5029 (0.220)	0.2301 (0.292)	1.0494 (0.393)	0.3971 (0.295)
$\ln X_4$	0.2992 (0.152)	0.6284 (0.420)	0.4836 (0.116)	0.3028 (0.154)	0.5682 (0.207)	0.4394 (0.156)
$(\ln X_1)^2$	-3.3675 (1.334)	. .	-0.6710 (1.019)	-4.8030 (1.353)	-1.9925 (1.820)	-1.3504 (1.366)
$(\ln X_3)^2$	0.1812 0.095)	0.0412 (0.263)	0.1575 (0.073)	0.2871 (0.097)	0.1305 (0.130)	0.2230 (0.098)
$(\ln X_3)^2$	-0.3918 (0.188)	0.3550 (0.519)	-0.0962 (0.144)	-0.3155 (0.191)	-0.0832 (0.257)	-0.2872 (0.193)
$(\ln X_4)^2$	-0.0189 (0.033)	0.0655 (0.091)	0.0029 (0.025)	-0.0280 (0.033)	0.0101 (0.045)	-0.0301 (0.034)
$\ln X_1.\ln X_2$	-0.5692 (0.587)	0.6863 (1.571)	-0.6105 (0.449)	-0.8591 (0.596)	0.0808 (0.801)	-0.5650 (0.601)
$\ln X_1.\ln X_3$	2.4390 (0.842)	0.2132 (2.332)	1.0273 (0.643)	2.0693 (0.854)	1.6578 (1.149)	1.9479 (0.862)
$\ln X_1.\ln X_4$	-0.3797 (0.327)	0.3026 (0.906)	-0.2264 (0.250)	0.1221 (0.332)	-0.3844 (0.447)	-0.5359 (0.335)
$\ln X_2.\ln X_3$	-0.9477 (0.192)	-0.5402 (0.532)	-0.3208 (0.147)	-1.0425 (0.195)	-0.4029 (0.263)	-0.4502 (0.197)
$\ln X_2.\ln X_4$	0.1606 (0.159)	0.0875 (0.441)	0.2679 (0.122)	0.1643 (0.162)	0.1426 (0.217)	0.1430 (0.163)
$\ln X_3.\ln X_4$	0.1559 (0.141)	0.0526 (0.390)	0.1437 (0.107)	0.1966 (0.143)	0.2950 (0.192)	0.2961 (0.144)
$\overline{R}^2$	0.9510	0.6574	0.8581	0.9521	0.7868	0.8866

Table 8. (continued)

Independent Variables, log form[b]	Dependent Variable, Yg[a]					
	Y13	Y14	Y15	Y16	Y17	Y18
Intercept	6.1226 (0.121)	6.2857 (0.139)	6.0585 (0.149)	5.0848 (0.168)	6.4553 (0.279)	6.2776 (0.126)
$\ln X_1$	-0.8239 (1.065)	0.0546 (1.222)	0.5650 (1.309)	0.8278 (0.857)	-0.0486 (1.430)	0.0019 (1.107)
$\ln X_2$	0.1745 (0.177)	0.2874 (0.203)	0.0763 (0.218)	0.1649 (0.236)	0.6082 (0.394)	0.2533 (0.184)
$\ln X_3$	0.6617 (0.349)	0.8685 (0.400)	1.0417 (0.429)	0.7491 (0.485)	1.2526 (0.809)	0.8791 (0.363)
$\ln X_4$	0.4593 (0.184)	0.4823 (0.2110	0.5309 (0.226)	0.4349 (0.259)	0.5647 (0.431)	0.5273 (0.191)
$(\ln X_1)^2$	-1.6031 (1.615)	-1.1551 (1.853)	-0.2001 (1.984)	. .	. .	-0.6561 (1.679)
$(\ln X_2)^2$	0.1365 (0.115)	0.0962 (0.132)	0.0646 (0.142)	0.0902 (0.162)	0.0837 (0.270)	0.1632 (0.120)
$(\ln X_3)^2$	-0.0040 (0.228)	-0.1041 (0.2610	-0.0698 (0.280)	-0.1598 (0.320)	0.2090 (0.5330	-0.0893 (0.237)
$(\ln X_4)^2$	-0.0096 (0.040)	0.0160 90.045)	0.0119 (0.049)	0.0155 (0.056)	0.0328 (0.093)	0.0272 (0.0410
$\ln X_1 . \ln X_2$	-0.3679 (0.711)	0.0648 (0.816)	-0.1333 (0.874)	-0.2301 (0.967)	0.5294 (1.613)	-0.7229 (0.739)
$\ln X_1 . \ln X_3$	0.7354 (1.019)	1.7590 (1.169)	1.8057 (1.252)	1.8840 (1.435)	0.2481 (2.394)	1.5590 (1.059)
$\ln X_1 . \ln X_4$	0.1815 (0.396)	-0.1682 (0.455)	-0.5652 (0.487)	-0.4549 (0.558)	-0.0978 (0.930)	-0.2740 (0.412)
$\ln X_2 . \ln X_3$	-1.0417 (0.233)	-0.7264 (0.267)	-0.5544 (0.286)	-0.3927 (0.327)	-0.8057 (0.546)	-0.3742 (0.242)
$\ln X_2 . \ln X_4$	0.2367 (0.193)	0.0743 (0.2210	0.2053 (0.237)	0.0901 (0.271)	0.2822 (0.453)	0.0875 (0.200)
$\ln X_3 . \ln X_4$	0.1036 (0.170)	0.2198 (0.195)	0.2770 (0.209)	0.2524 (0.240)	0.1812 (0.400)	0.2099 (0.177)
$\overline{R}^2$	0.9423	0.8797	0.7417	0.6087	0.6519	0.8050

Table 8. (continued)

Independent Variables, log form[b]	Dependent Variable, Yg[a]			
	Y19	Y20	Y21	Y22
Intercept	6.2608 (0.181)	6.3729 (0.212)	5.6690 (0.1120	6.0302 (0.210)
ln X_1	0.0115 (0.928)	0.9058 (1.863)	-0.4120 (0.984)	2.3045 (1.847)
ln X_2	0.5272 (0.256)	0.2300 (0.310)	-0.0138 (0.1640	0.4383 (0.307)
ln X_3	1.2022 (0.526)	1.2983 (0.611)	-0.0111 (0.322)	0.7152 (0.605)
ln X_4	0.4161 (0.280)	0.6591 (0.322)	0.5283 (0.170)	0.6578 (0.319)
$(\ln X_1)^2$	. .	0.7644 (2.825)	-1.7957 (1.492)	1.6963 (2.800)
$(\ln X_2)^2$	0.1356 (0.175)	0.0993 (0.202)	0.2033 (0.106)	0.0091 (0.200)
$(\ln X_3)^2$	0.1591 (0.346)	0.0258 (0.398)	-0.5027 (0.210)	-0.1896 (0.395)
$(\ln X_4)^2$	-0.0064 (0.060)	0.0628 (0.069)	-0.0082 (0.037)	-0.0208 (0.069)
$\ln X_1.\ln X_2$	0.0056 (1.0470	0.0559 (1.244)	-1.0286 (0.657)	1.0842 (1.233)
$\ln X_1.\ln X_3$	0.4984 (1.554)	1.4456 (1.783)	2.7875 (0.941)	2.0861 (1.767)
$\ln X_1.\ln X_4$	-0.0021 (0.604)	-0.1541 (0.693)	-0.4636 (0.366)	-0.6515 (0.687)
$\ln X_2.\ln X_3$	0.4984 (0.354)	1.4456 (0.408)	2.7875 (0.215)	2.0861 (0.404)
$\ln X_2.\ln X_4$	0.2498 (0.294)	0.0150 (0.337)	0.2135 (0.178)	0.4011 (0.334)
$\ln X_3.\ln X_4$	0.1740 (0.260)	0.1707 (0.298)	0.3359 (0.157)	0.3557 (0.295)
$\bar{R}^2$	0.7487	0.8707	0.9392	0.8681

[a] Yg is genotype numbers. See Appendix A for description of genotypes.

[b] X_1 through X_4 are relative water availabilities, RAW_i for the four growth phases.

Table 9. Theil Coefficients for Predicted Yields from Final Translog Models

GNO[a]	U	U^M	U^S	U^C
1	0.0974	0.0019	0.0399	0.9582
2	0.1101	0.0102	0.0068	0.9830
3	0.1074	0.0078	0.0005	0.9917
4	0.0857	0.0140	0.0061	0.9799
5	0.1152	0.0138	0.0609	0.9253
6	0.0995	0.0164	0.0003	0.9832
7	0.1159	0.0065	0.0583	0.9352
8	0.1383	0.0051	0.0210	0.9739
9	0.0779	0.0182	0.0284	0.9534
10	0.0988	0.0033	0.0062	0.9905
11	0.1013	0.0237	0.0017	0.9746
12	0.0965	0.0164	0.0316	0.9519
13	0.0952	0.0108	0.0101	0.9791
14	0.1035	0.0104	0.0005	0.9892
15	0.1237	0.0253	0.0263	0.9484
16	0.1423	0.0282	0.0394	0.9324
17	0.1234	0.0083	0.0361	0.9556
18	0.0968	0.0169	0.0034	0.9797
19	0.1008	0.0222	0.0054	0.9724
20	0.1178	0.0129	0.0110	0.9761
21	0.1136	0.0011	0.0406	0.9584
22	0.1266	0.0044	0.0171	0.9785

Note: U estimated from equation (29), U^M from equation (32), U^S from equation (33) and $U^C = 1 - U^M - U^S$.

[a]Numbers refer to genotypes. See Appendix A for descriptions.

The marginal response to RAW_3 is reduced, and for some genotypes is negative, while the marginal response to RAW_4 increases for some genotypes and is decreased for others. The marginal product of RAW_1 varies, but many genotypes exhibit a large positive marginal response to RAW in the early phase. Holding values of RAW in all other growth phases constant at their maximum value, marginal response to RAW_1 is positive but declining when RAW_1 is at its minimum value, and becomes negative as RAW_1 approaches its maximum value of 1.

Second, Nageswara Rao et al. (1985) noted that the response to a late season drought was modified according to whether or not the early phase was droughted. RAW_i were set equal to 0.5 for all i. This represents the situation in which RAW_1 is at its minimum value, and RAW_3 and RAW_4 are close to their mean values. By comparing this situation to the first scenario in which RAW in all phases were set equal to their mean values, we can examine how an increase in stress in the early phase affects the response to late season stress. Again, results vary across genotype. Marginal response to stress in phase 3 is reduced for most genotypes (except GNO8 and GNO17), while in phase 4 it is increased for some genotypes and decreased for others.

Simulation of Yields in Three Sample Locations

The translog cannot be used to predict yields when any $X_i = 0$, therefore, in predicting yields in the three study sites, zero values of X_i were set equal to 0.01, the minimum amount of water applied in the highest intensity of the long term drought pattern, P7, in which the crop received virtually no water for the extent of the experiment.

The exponential of the intercept in Table 8 gives the predicted yields when all $X_i = 1$ (no drought). These compare favorably with maximum yields recorded in the experiment and in other irrigated post rainy season trials at ICRISAT, but they are high compared with the yields normally obtained in the rainy season. The measure of RAW developed in this study does not adequately reflect the differences in meteorological conditions between the two seasons. The model is not intended to predict what yields would have actually been in any given year in any given location; it is argued, however, that the relative ordering of genotypes, in terms of yield, is maintained. The simulated yields are not, therefore, to be regarded as representing actual yields, but should be regarded as indices of genotype performance.

Summary statistics for the distributions of simulated yields in Hyderabad, Anantapur, and Gujarat are given in Tables 10, 11 and 12, respectively. For Hyderabad and Anantapur, these statistics include years in which it was determined that rainfall was insufficient to allow planting of groundnuts; these years are included in the analysis with simulated yields set at zero. There are a number of years in Hyderabad in which no drought was recorded, $RAW = 1$ in all growth phases. This is reflected in the values of the median and higher percentiles, in Table 10, which are close to, or equal, the maximum yields and, consequently, the distributions of yields in Hyderabad are highly skewed to the left, observations being concentrated at higher values of yield. The CDFs will have long lefthand tails then turn sharply upward as yields approach their maximum value.

A similar situation occurs in Gujarat, where there were a number of years in which no drought was recorded. However, due to the short rainy season in Gujarat, there were also a number of years in which late season drought severely reduced yields, as indicated by the minimum yields in Table 12. The yield distributions are also negatively skewed, but not to the extent of those in Hyderabad. Note that the lower percentiles in Gujarat are very low, while the median and upper percentiles approach the maximum yields. Consequently, yields in Gujarat tend toward a bimodal type of distribution, observations being concentrated in the upper and lower ranges of yield. The CDFs will have a steplike appearance, with a gradual increase in the lower tail, a long flat section over the mid range of yields, rising sharply as they approach maximum yields.

Anantapur not only experiences droughts more often than either Hyderabad or Gujarat, but is also prone to long term droughts with the result that minimum yields for many genotypes are negligible. The greater frequency of droughts is demonstrated by the values of the lower percentiles and the median. Distributions in Anantapur are positively skewed, observations being concentrated below the mean. The degree of skewness is also more variable, across genotypes, than in the other two sites. CDFs will have a smoother, more familiar S-shaped appearance.

Subsequent to the analysis reported here, in an effort to reduce the high incidence of years (15 out of 74) in Anantapur in which there was insufficient rainfall on which to plant groundnuts, decision rules for planting were relaxed and planting dates were determined for a further 12 years. RAW_i were estimated and yields for these years were simulated. Reducing the rainfall requirement in the planting rules means that in many years, soil moisture at date of planting was very low (< 20mm. for 9 of the 12 years). This, associated

Table 10. Summary Statistics for Simulated Yields (kg./ha. by genotype), Hyderabad

GNO[a]	Maximum	Minimum	Mean	Std. Dev.	Skewness	Median
1	4,470	0	3,612	1,189	-1.5	4,245
2	5,331	0	4,272	1,401	-1.4	4,999
3	5,522	0	4,431	1,444	-1.3	5,146
4	6,348	0	5,106	1,650	-1.4	5,905
5	4,816	0	3,631	1,002	-1.6	4,147
6	5,982	0	4,610	1,465	-1.3	5,304
7	3,061	0	2,444	606	-2.6	2,720
8	6,821	0	5,171	2,046	-1.0	6,136
9	3,539	0	3,004	817	-2.1	3,391
10	3,625	0	2,767	757	-2.1	3,128
11	5,781	0	4,545	1,597	-1.2	5,339
12	3,925	0	3,255	963	-1.8	3,723
13	4,834	0	3,843	1,111	-1.9	4,468
14	5,368	0	4,347	1,381	-1.4	5,065
15	4,278	0	3,451	1,122	-1.4	4,040
16	3,319	0	2,737	804	-1.6	3,160
17	6,361	0	5,013	1,763	-1.2	5,882
18	5,324	0	4,354	1,335	-1.5	5,036
19	5,237	0	4,169	1,409	-1.3	4,805
20	5,858	0	4,577	1,634	-1.1	5,400
21	3,597	0	2,568	675	-2.5	2,897
22	4,158	0	3,257	1,158	-1.2	3,917

GNO	Percentiles 5^{th}	10^{th}	25^{th}	75^{th}	90^{th}	95^{th}
1	474	2,024	2,866	4,470	4,470	4,470
2	819	2,416	3,235	5,331	5,331	5,331
3	1,024	2,422	3,384	5,522	5,522	5,522
4	1,130	2,811	3,914	6,348	6,348	6,348
5	1,790	2,253	2,888	4,332	4,332	4,332
6	1,624	2,530	3,378	5,696	5,696	5,696
7	761	1,762	2,292	2,736	2,736	2,900
8	1,113	2,032	3,092	6,821	6,821	6,821
9	781	2,071	2,554	3,539	3,539	3,539
10	674	1,871	2,469	3,177	3,177	3,399
11	825	2,412	3,022	5,781	5,781	5,781
12	650	2,103	2,729	3,925	3,925	3,925
13	819	2,474	3,273	4,560	4,560	4,560
14	1,073	2,506	3,198	5,368	5,368	5,368
15	738	2,036	2,549	4,278	4,278	4,278
16	842	1,677	2,183	3,319	3,319	3,319
17	999	2,479	3,440	6,361	6,361	6,361
18	1,098	2,625	3,375	5,324	5,324	5,324
19	834	2,121	2,947	5,237	5,237	5,237
20	1,121	2,142	3,291	5,858	5,858	5,858
21	596	1,736	2,432	2,897	2,899	2,930
22	352	1,729	2,369	4,158	4,158	4,158

[a]See Appendix A for description of genotypes.

Table 11. Summary Statistics for Simulated Yields (kg./ha. by genotype), Anantapur

GNO[a]	Maximum	Minimum	Mean	Std. Dev.	Skewness	Median
1	4,189	0	964	1,141	1.0	475
2	4,957	0	1,034	1,229	1.2	555
3	5,179	0	1,170	1,355	1.2	646
4	6,083	0	1,142	1,260	1.4	752
5	5,270	0	1,156	1,024	1.0	1,173
6	5,779	0	1,330	1,428	1.2	819
7	3,473	0	843	913	1.1	770
8	6,699	0	1,192	1,631	1.7	433
9	3,730	0	975	843	0.7	921
10	4,060	0	742	928	1.7	515
11	5,761	0	935	993	2.0	853
12	3,777	0	727	778	1.5	530
13	5,062	0	967	1,172	1.3	457
14	4,999	0	1,064	1,115	1.1	866
15	3,726	0	971	901	0.7	885
16	2,752	0	813	673	0.4	868
17	6,431	0	1,382	1,715	1.3	700
18	5,146	0	1,059	1,048	1.3	880
19	5,223	0	1,055	1,214	1.2	587
20	5,570	0	1,012	1,345	1.5	449
21	3,725	0	827	942	1.3	678
22	4,065	0	764	941	1.3	418

	Percentiles					
GNO	5^{th} [b]	10^{th} [b]	25^{th}	75^{th}	90^{th}	95^{th}
1	0	0	1	1,740	2,749	3,143
2	0	0	3	1,934	2,886	3,247
3	0	0	43	1,863	3,419	4,123
4	0	0	118	1,814	3,033	3,798
5	0	0	114	1,904	2,345	2,746
6	0	0	181	2,147	3,545	4,403
7	0	0	6	1,322	2,311	2,916
8	0	0	41	1,858	3,959	5,083
9	0	0	164	1,485	2,246	2,514
10	0	0	5	1,163	2,009	3,010
11	0	0	92	1,318	2,122	2,769
12	0	0	53	1,009	1845	2,491
13	0	0	7	1,735	2888	3,220
14	0	0	19	1,648	2681	3,308
15	0	0	49	1,608	2269	2,702
16	0	0	110	1,279	1682	2,043
17	0	0	18	2,260	4409	5,337
18	0	0	127	1,616	2625	3,088
19	0	0	26	1,876	2889	3,390
20	0	0	2	1,518	3288	4,089
21	0	0	6	1,179	2531	2,869
22	0	0	1	1,299	2232	2,702

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]In 15 years (25% of the sample), no groundnuts were planted and simulated yields = 0.

Table 12. Summary Statistics for Simulated Yields (kg./ha. by genotype), Gujarat

GNO[a]	Maximum	Minimum	Mean	Std. Dev.	Skewness	Median
1	5,783	8	2,760	2,040	-0.3	4,466
2	6,618	8	2,760	2,040	-0.3	4,466
2	6,618	6	3,405	2,304	-0.4	5,319
3	5,802	203	3,535	2,320	-0.4	5,518
4	6,348	213	3,983	2,605	-0.4	5,459
5	4,379	210	3,082	1,536	-0.7	4,089
6	5,696	171	3,679	2,243	-0.4	4,935
7	3,195	0	1,889	1,065	-0.5	2,736
8	6,821	83	4,171	2,895	-0.3	5,891
9	3,549	48	2,236	1,445	-0.4	3,064
10	3,283	0	2,142	1,258	-0.5	3,177
11	5,799	59	3,540	2,424	-0.3	4,842
12	3,928	6	2,452	1,628	-0.4	3,438
13	5,420	22	2,917	1,986	-0.4	4,560
14	5,368	14	3,458	2,154	-0.4	4,946
15	4,278	51	2,700	1,760	-0.4	3,909
16	3,319	21	2,200	1,233	-0.4	2,887
17	6,363	21	4,035	2,646	-0.4	6,327
18	5,324	38	3,373	2,122	-0.3	4,529
19	5,237	76	3,297	2,185	-0.4	4,716
20	7,208	10	3,728	2,503	-0.3	5,839
21	2,906	0	1,909	1,158	-0.5	2,826
22	4,240	4	2,539	1,852	-0.4	4,124

	Percentiles					
GNO	5^{th}	10^{th}	25^{th}	75^{th}	90^{th}	95^{th}
1	94	197	294	4,470	4,470	4,930
2	257	492	590	5,331	5,331	5,781
3	342	499	678	5,522	5,522	5,620
4	430	652	825	6,348	6,348	6,348
5	280	657	1,794	4,332	4,332	4,348
6	346	615	1,122	5,696	5,696	5,696
7	294	468	584	2,736	2,736	2,916
8	209	349	818	6,821	6,821	6,821
9	260	413	433	3,539	3,539	3,542
10	291	489	568	3,177	3,177	3,247
11	362	545	635	5,781	5,781	5,787
12	187	382	517	3,925	3,925	3,926
13	206	330	416	4,560	4,560	4,879
14	340	721	854	5,368	5,368	5,368
15	323	492	551	4,278	4,278	4,278
16	340	657	777	3,319	3,319	3,319
17	395	743	819	6,361	6,361	6,362
18	415	758	850	5,324	5,324	5,324
19	327	506	581	5,237	5,237	5,237
20	168	494	854	5,858	5,858	6,330
21	148	363	543	2,897	2,897	2,900
22	89	180	254	4,158	4,158	4,187

[a]See Appendix A for description of genotypes.

with the fact that in many of these years planting was followed by a long dry spell, leads to very low values of RAW_1. Furthermore, planting rules were extended into the first few days of September. Such late planting dates extend the growing season into the post rainy season, increasing the incidence of late season drought.

In a number of these years, simulated yields for some genotypes were highly inflated when others approached zero. The problems are two fold, one associated with the design of the experiment, the other with the estimated response functions.

In the experiment, the plots were uniformly irrigated for the first 29 days to ensure a good crop stand; consequently, the carryover of soil moisture ensured that RAW_1 is never less than 0.5, even in the highest intensity of early drought treatments.

Second, when RAW_1 becomes very small, or patterns of drought occur that are not represented in the trial, the estimated translog response functions are not always well-behaved. Highly inflated yields are predicted for some genotypes when $RAW_1 \rightarrow 0$ ($RAW_1 < 0.1$ for 6 of the 12 years), especially when associated with extreme drought in the second phase ($RAW_2 \rightarrow 0$), or with no drought in the fourth phase ($RAW_4 \rightarrow 1$), a pattern that is not represented in the trial. However, even if the response functions had been well behaved, it is likely that yields in most of these years would have been minimal, with the result that including these years would merely extend the lower tails of the simulated yield distributions.

For four of the 12 years, the translog function predicted reasonable, low yields for all genotypes. Including these observations in the simulated yield distributions for Anantapur increased mean yields by, on average, 2.1%, and reduced standard deviations by, on average, 1.6%. In subsequent sections on risk analysis, sensitivity of results to the inclusion of these observations is noted, where appropriate.

VII - RISK ANALYSIS

Using these simulated groundnut yields, it is now possible to assess how a selection of groundnut genotypes perform in diverse and risky environments. It was shown in section *II* that yield distributions of different genotypes may exhibit cross-overs in lower yielding environments, and that there may exist a trade-off between yield potential under favorable conditions and the stability, or conversely the variability, of performance over a range of environmental conditions. One purpose of the analysis in this section is to determine whether such cross-overs exist.

One concern is with the empirical identification of risk efficient genotypes for the three sample locations : Hyderabad, Anantapur, and Gujarat. Another is with the methodological issues associated with the different analytical approaches. While a trial such as that conducted by ICRISAT may approximate the range of environmental conditions occurring in a region, genotypes identified as stable or widely adaptable on the basis of the trial data alone will not necessarily be the preferred genotype in any given location within the region. By comparing results from the three sample locations with those from the trial, one can see how recommendations based on the trial results alone would compare with those obtained by incorporating location specific meteorological information.

Results from the approach developed by Finlay and Wilkinson (1963) and Eberhart and Russell (1966) using the actual yields from the trial is an initial point for comparison. The same kind of analysis can be conducted on the simulated yields in each sample location.

The results from the alternative approaches to risk analysis discussed in section *III* are then presented and compared to this 'base' approach both across and within locations.

Stability Analysis

Actual Experimental Yields

Recalling that each drought pattern in the trial included a control non-stressed treatment (intensity 1), yields were averaged over these plots so that results would not be biased toward higher yielding environments. Following equation (1), actual yields (in kg./ha.) for each genotype, in each treatment were regressed on the environmental index, the mean yield over all genotypes in each treatment. Genotype mean yields, and the regression coefficients a_g and b_g, are presented in Appendix C, ranked by mean yield. The uncorrected MSE for each genotype is taken as a proxy for $s^2{}_d$ and included in the appendix table for completeness; the two measures of stability, b_g and $s^2{}_d$, are not complementary, in fact they are negatively correlated (r=-0.42), and lead to completely different rankings, in terms of stability, of genotypes.

Following Finlay and Wilkinson (1963), the stability parameters, b_g, are plotted against genotype mean yields ($\overline{Y}_g.$) in Figure 7. The plot is divided into four quadrants; the vertical dividing line is the overall average yield of all genotypes over all treatments, the horizontal line represents average stability, when $b_g = 1$. If a decision maker wishes to maximize stability (minimize b_g) and maximize mean yield, then we can identify a number of genotypes that can be eliminated from consideration. In a manner similar to EV analysis, genotypes are discarded if there exists another genotype that lies to the right and below it in mean-stability (Eb_g) space. In this manner the "stability efficient" set of genotypes can be identified and contains GNO4, GNO9, GNO16 and GNO17.[5]

Plotting the regression lines from equation (1) demonstrates the relative performance of genotypes over the range of environments (treatments) in the trial (Figure 8). Of the genotypes from the stability efficient set, GNO4 and GNO17 are high yielding, but differ in their degree of stability, while GNO9 and GNO16 are lower yielding and more stable. GNO19 is included as an average performer both in terms of yield and stability. It can be seen from Figure 8 that cross-overs in yield performance do occur in environments (treatments) with lower yielding potential. The importance of these cross-overs depends on the likelihood of such low-yielding environments occurring. If the probability of such events is low, then GNO4 would be the preferred genotype.

In assessing the performance of genotypes in this way, one must remember that the environmental index is a function of the environments contained in the trial; to assess the importance of cross-overs in yield, and the trade-off between yield and stability, the same type of analysis is conducted on the simulated yield data for the three sample sites. To account for the error structure associated with the estimated response functions used in the simulation of yields, results from each of the three sites should be compared with those from the stability analysis for the predicted experimental yields from the estimated response functions.

[5] The "stability" efficient sets include only those genotypes where the linear line segments connecting the points of minimum instability for a given mean yield forms a convex point set. In this way, using a convex combination of efficient genotypes leads to more stable yield for a given mean than any single inefficient genotype. In this case, the efficiency locus is only approximate in that the covariance between yields of the individual genotypes is not considered explicitly.

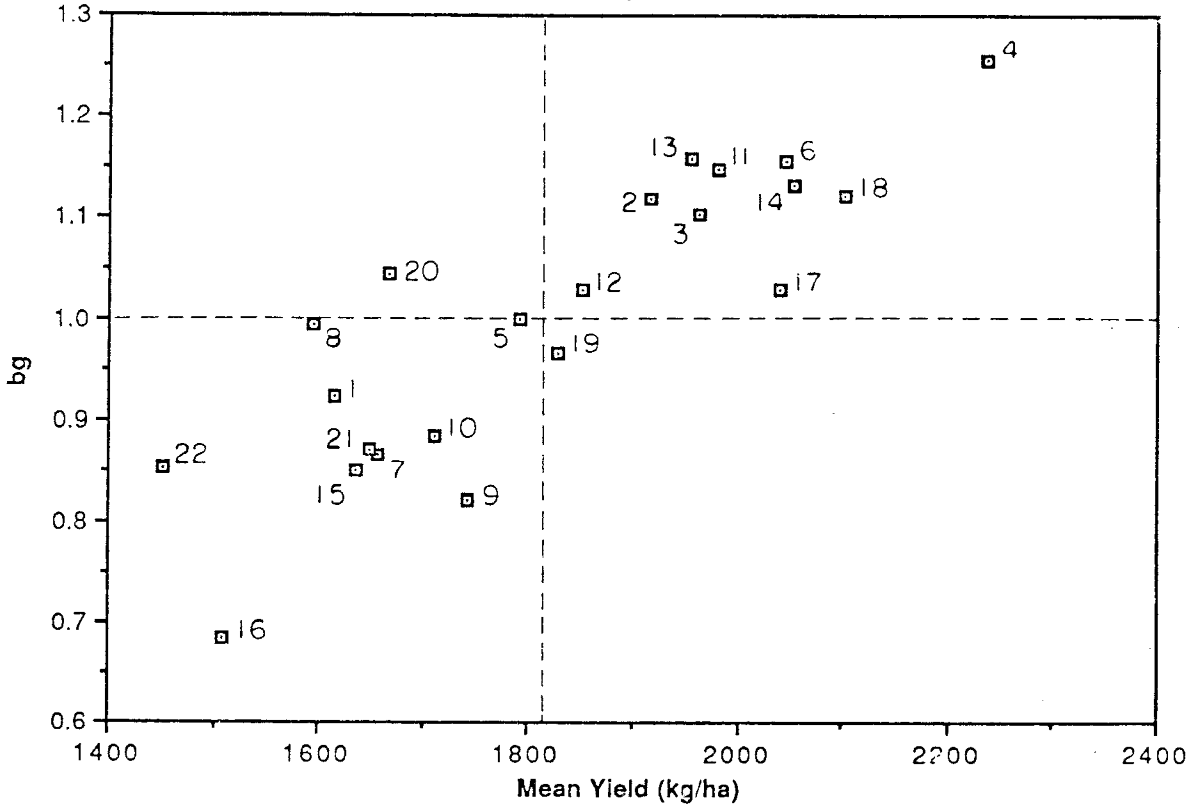

Figure 7. Plot of stability parameters against mean yields. Actual yields from trial

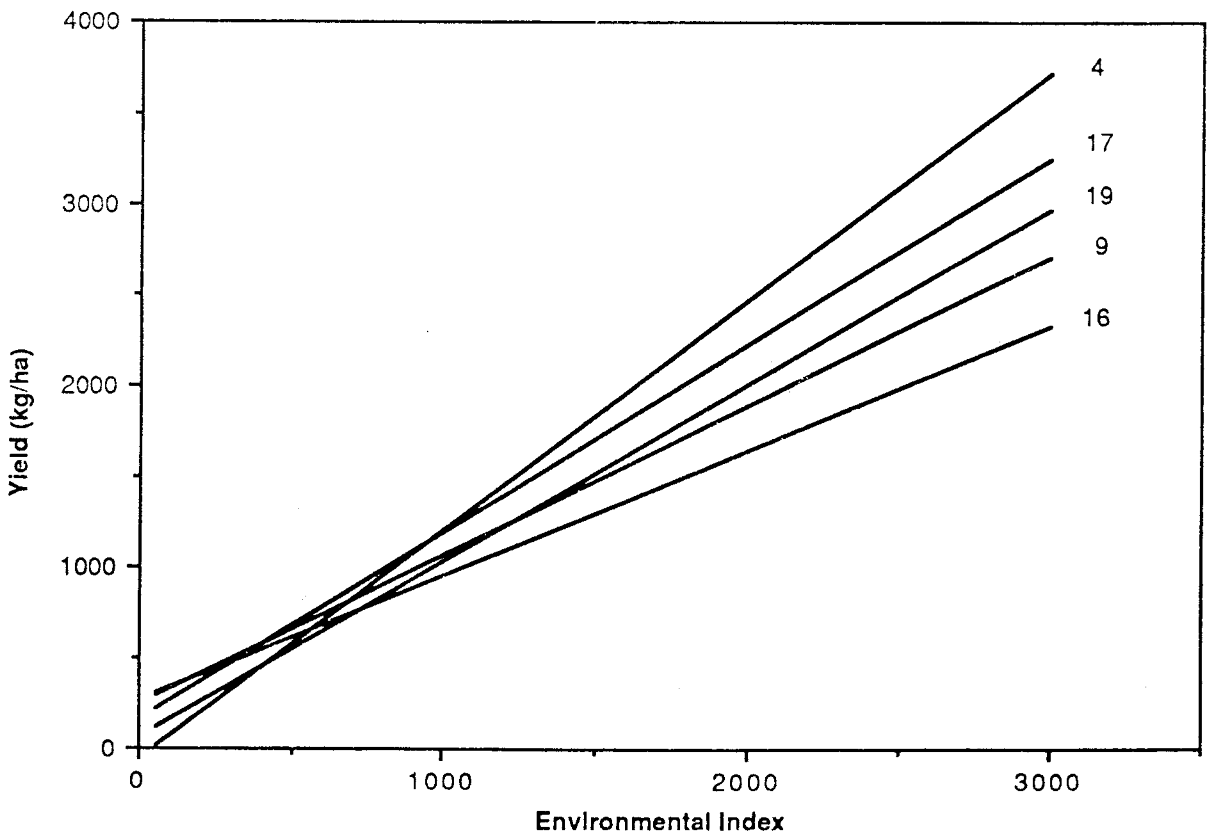

Figure 8. Genotype-environment interaction of selected genotypes. Actual yields from trial

Predicted Experimental Yields

The slope coefficients from the regression of predicted yields on the experimental environmental index are also presented in Appendix C, along with the mean yield and uncorrected MSE for each genotype. Mean yields are plotted against b_g in Figure 9. The partition of genotypes into the different quadrants is essentially the same as the actual experimental yields with a few exceptions. Most noticeable is the change in position of GNO17 which exhibits a higher degree of instability and is no longer stability efficient; GNO18 replaces GNO17 in the stability efficient set. The regression lines for selected genotypes are plotted in Figure 10 and are easily compared with those in Figure 8. Again, cross-over points correspond to the highest intensities of drought in late-season or long-term drought patterns in the trial.

On the basis of these results, one would be tempted to recommend GNO4 for locations that infrequently experience drought, and GNO18 or GNO9 for locations with a greater likelihood of late-season or long-term droughts. Whether such recommendations are sound can be examined by conducting the same type of analysis on the simulated yields for each of the sample locations.

Simulated Yields: Hyderabad

It should be recalled from section *VI* that there are a number of years in Hyderabad and Anantapur in which there was insufficient moisture on which to plant a groundnut crop. For these years no growing season was defined, and yields are set to zero and included in the analysis.

Genotype yields in each year t, Y_{gt}, were regressed on an environmental index, the mean yield of all genotypes in each year $\overline{Y}_{.t}$. Mean yields, the regression coefficients, and uncorrected MSE are given in Appendix C. In 40% of the years in Hyderabad, no drought (as measured by RAW) occurred and therefore mean yields are much higher than in the trial. The stability parameters, b_g, are plotted against mean yields in Figure 11. The genotypes are equally divided between the north-east and south-west quadrants, and follow a more or less linear pattern, stability decreasing as mean yield increases. The stability efficient set of genotypes contains GNO8, GNO4, GNO5, and GNO7. A noticeable new entry is GNO8; its position relative to its position in the experimental results (Figure 9) may be attributed to the fact that the experiment did not cover the full range of environments occurring in Hyderabad as measured by RAW in the four growth stages. Within the experimental range, GNO8 does not perform particularly well, but when simulating yields outside this range, GNO8 responds well to increases in RAW in growth stages G3 and G4. Consequently, its mean yield in Hyderabad is comparable to that of GNO4, while its stability parameter is far higher due to its poorer performance in suboptimal conditions. GNO8, in this case, is an example of a highly unstable genotype (relative to the other genotypes included in the trial), specifically adapted to high yielding environments. A relatively small reduction in mean yield can be traded for a relatively large gain in stability by switching from GNO8 to GNO4.

The regression lines for a selection of genotypes are plotted in Figure 12. It can be seen that cross-overs occur among the high yielding, less stable genotypes GNO4 and GNO8. GNO4 outyields GNO8 when $\overline{Y}_{.t} \leq$ 3,636 kg./ha., which represents approximately one third of the 84 years in Hyderabad. GNO4 remains the dominant genotype, in terms of yield, until the environmental index drops to 1,437 kg./ha., when it is replaced by GNO5; such a low index represents only two of the 84 years in the sample.

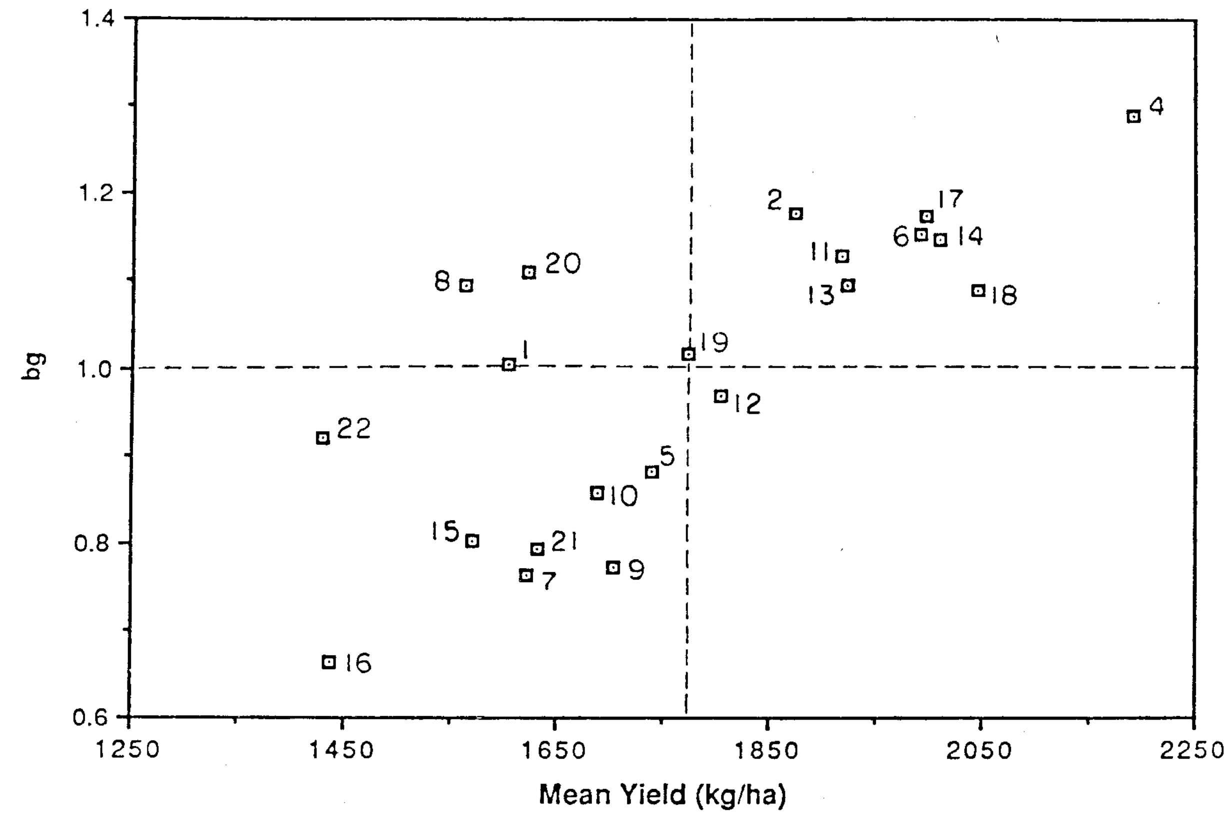

Figure 9. Plot of stability parameters against mean yields. Predicted yields from trial

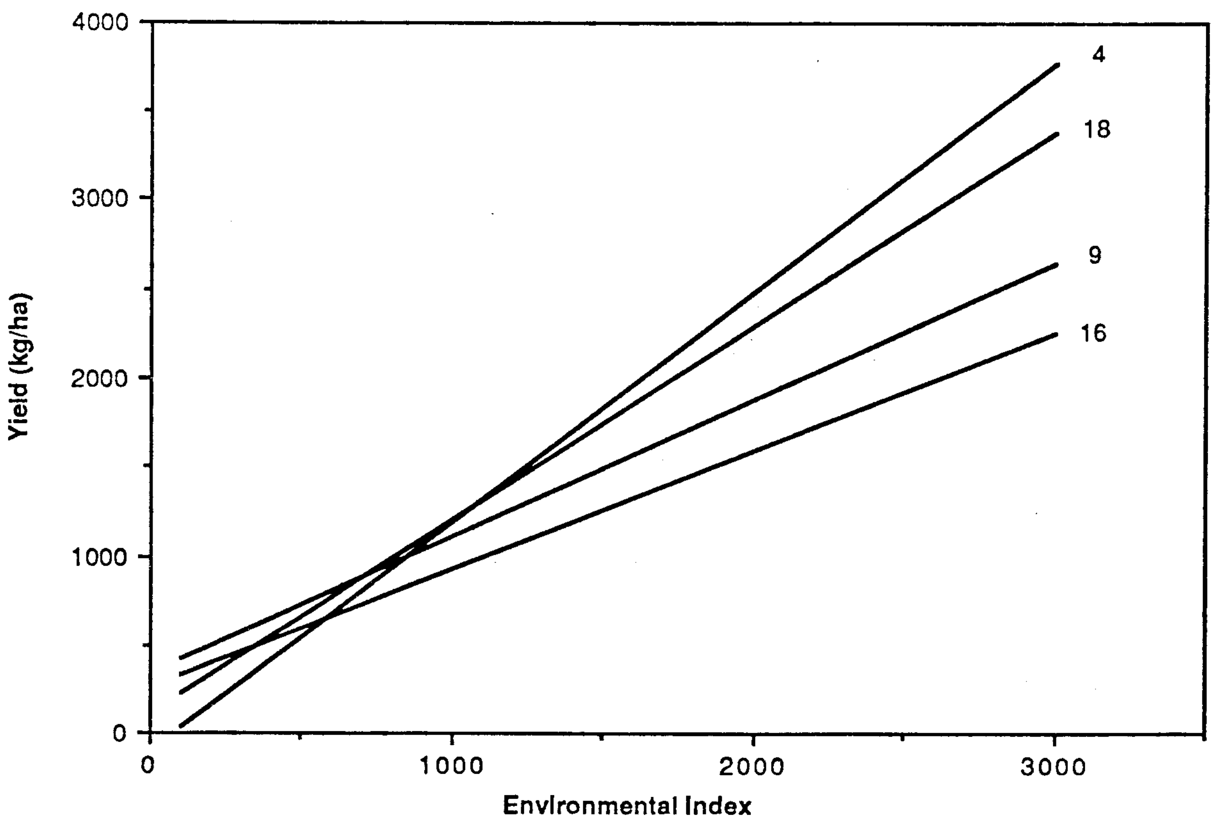

Figure 10. Genotype-environment interaction of selected genotypes. Predicted yields from trial

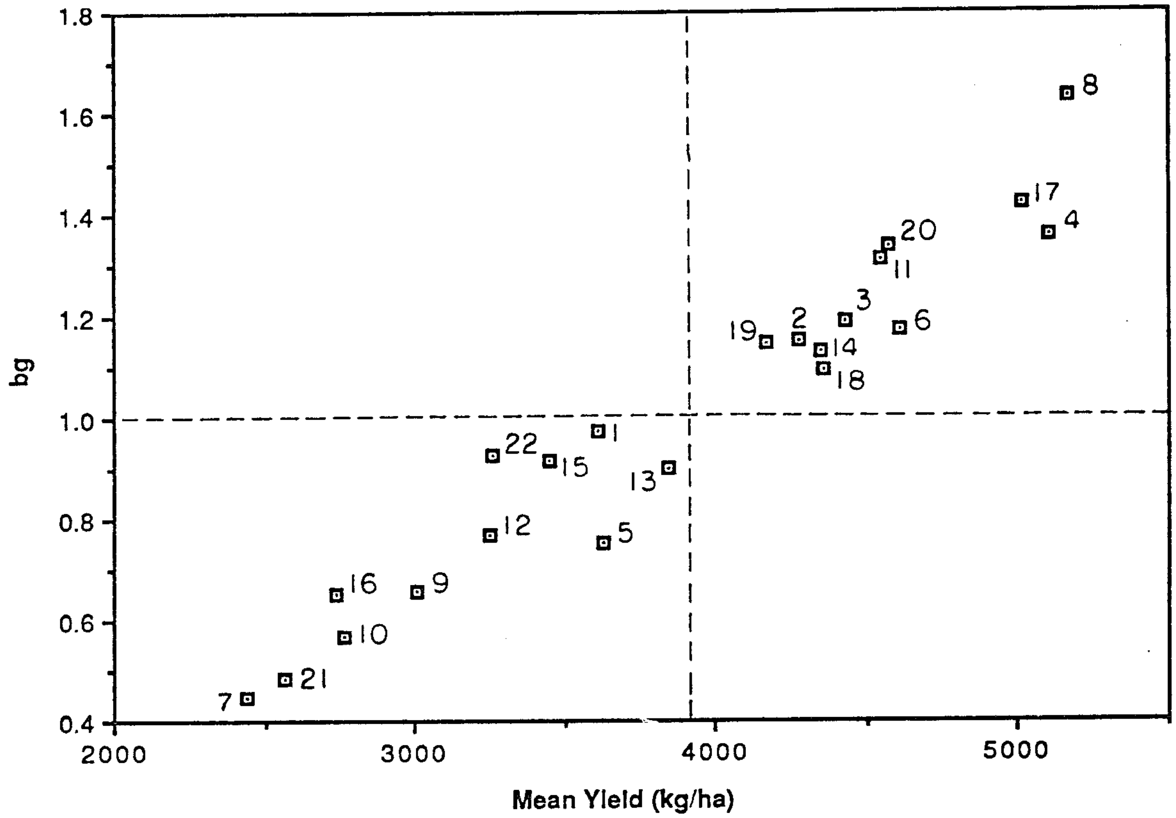

Figure 11. Plot of stability parameters against mean yields, Hyderabad

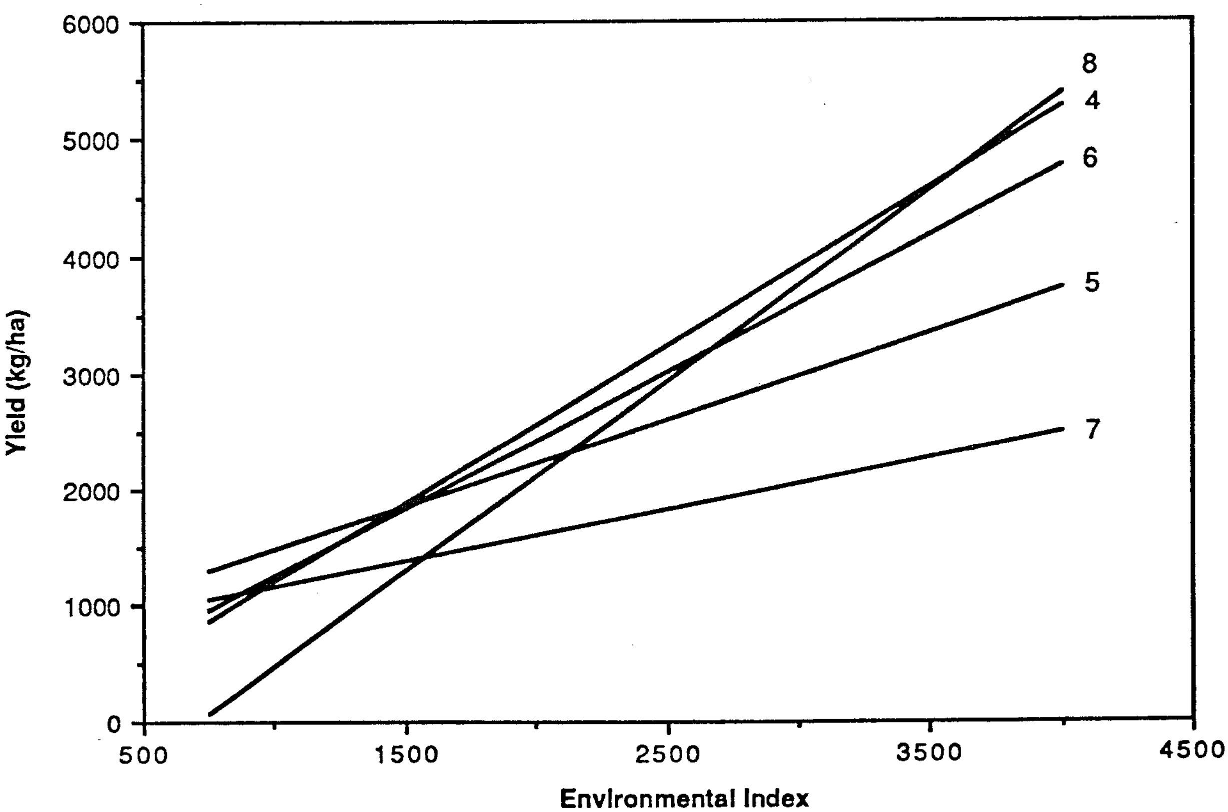

Figure 12. Genotype-environment interaction of selected genotypes, Hyderabad

Simulated Yields: Anantapur

In contrast to Hyderabad, meteorological conditions at Anantapur over the period 1911-1984 were highly variable with droughts occurring frequently in any of the four growth stages. Mean yields are low, as can be seen from the table in Appendix C. In plotting the stability parameters against mean yields (Figure 13), the genotypes are not so clearly partitioned into two quadrants. Most noticeable is the position of GNO5, which has a mean yield, in Anantapur, comparable to that of GNO4, the highest yielder in the trial, but also exhibits a high degree of stability. The inference is that GNO5 is well adapted to all environments occurring in Anantapur, but cannot take advantage of higher yielding environments such as those occurring in Hyderabad. The stability efficient set contains four genotypes in Anantapur: GNO17, GNO6, GNO5 and GNO16.

The regression lines for these genotypes are plotted in Figure 14. The cross-over in yield performance between GNO17 and GNO6 (at $\overline{Y}._t$ = 756 kg./ha.) occurs at almost the same point on the environmental scale as the cross-over between GNO6 and GNO5 (at $\overline{Y}._t$ = 706 kg./ha.). Just over one third of the years in Anantapur are represented by an environmental index below 706 kg./ha., and the problem of genotype selection in Anantapur, based on this type of stability analysis, reduces to a choice between two genotypes, GNO17 and GNO5. Inclusion of observations on yields for four more years, as discussed at the end of section VI, did not effect the results of the stability analysis for Anantapur.

Simulated Yields: Gujarat

Gujarat, with its short but intense rainy season, is prone to late season droughts. However, in 11 out of 32 years, no droughts occurred at all. Gujarat can, therefore, be viewed as an intermediate location. Like Hyderabad, there are a number of years in which yields achieve their predicted potential in non-stressed conditions, but there is also the possibility of severely reduced yields resulting from late season droughts. Consequently, as can be seen from Appendix C, mean yields are relatively high, but not as high as those in Hyderabad.

The plot of b_g against mean yields in Figure 15 is similar to that for Hyderabad, with the noticeable exception of GNO5 which approaches the position it holds in Anantapur. According to Finlay and Wilkinson's criteria, GNO5 appears to be a stable, widely adaptable genotype, but by Eberhart and Russell's second parameter of stability (MSE) it is highly unstable, having the highest MSE (Appendix C).

From Figure 16, a situation similar to that in Anantapur is seen to exist. While a cross-over in yield between GNO8 and GNO17 occurs at $\overline{Y}._t$ = 1,950 kg./ha., there is very little advantage to be gained in terms of yield, in switching from GNO8 to GNO17, before another cross-over occurs further down the environmental scale, between GNO5 and GNO17 at $\overline{Y}._t$ = 1,552 kg./ha. The 34 years included in the sample can be clearly divided into two groups, those with no, or at least very low intensities of, drought and those with severe late-season drought; 34% of years have an environmental index below 1,552 kg./ha., while the remainder all have an environmental index above 1,950 kg./ha., the point of cross-over between GNO8 and GNO17. The choice of genotype then reduces to a choice between GNO8 and GNO5. If we also take account of the fact that even in the highest yielding environments, the yield advantage of GNO8 over GNO17 amounts to about 250 kg./ha., then the choice is similar to that in Anantapur, between GNO17 and GNO5.

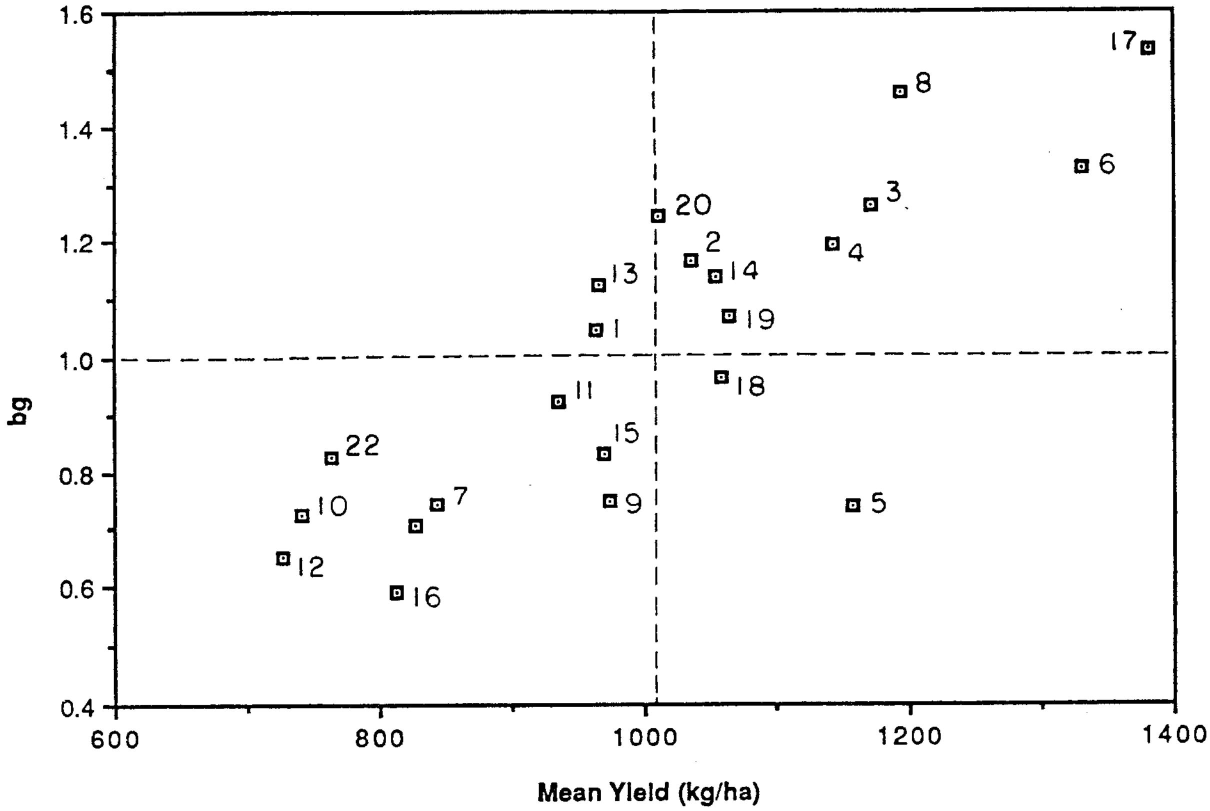

Figure 13. Plot of stability parameters against mean yields, Anantapur

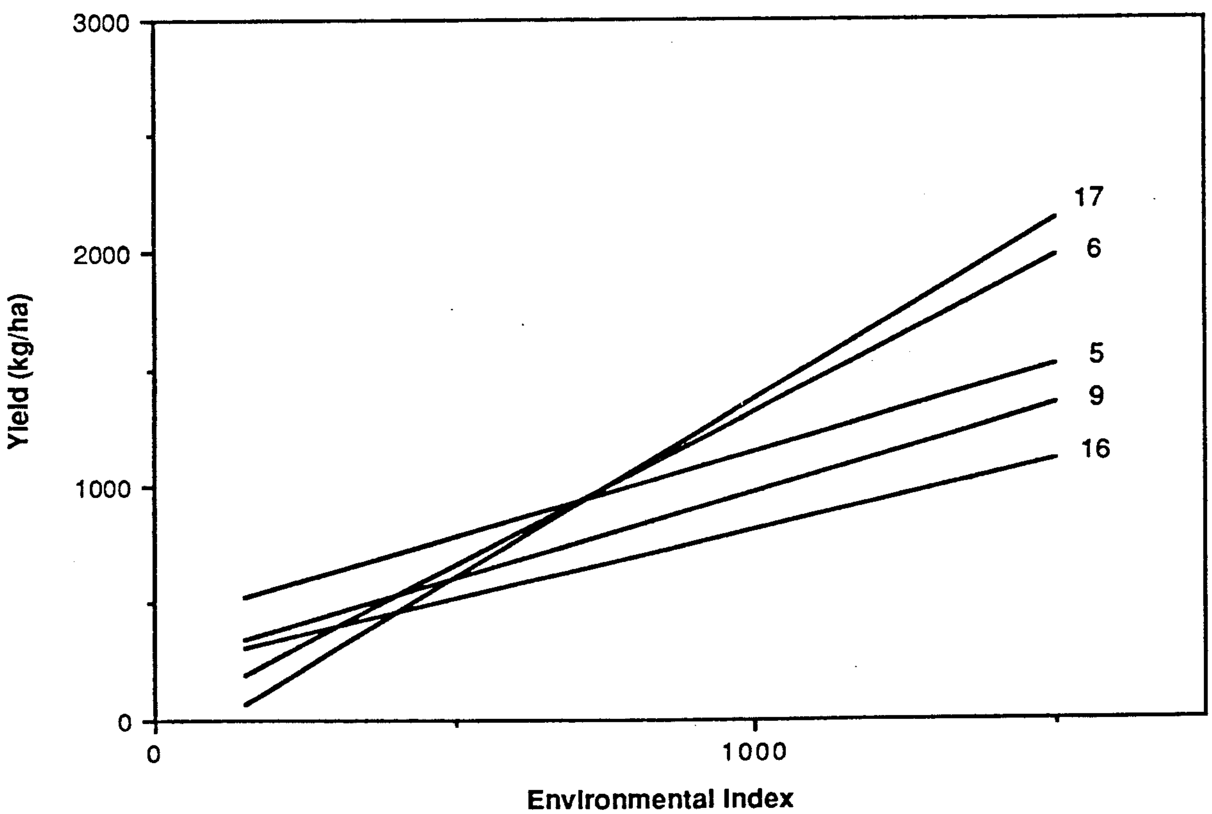

Figure 14. Genotype-environment interaction of selected genotypes, Anantapur

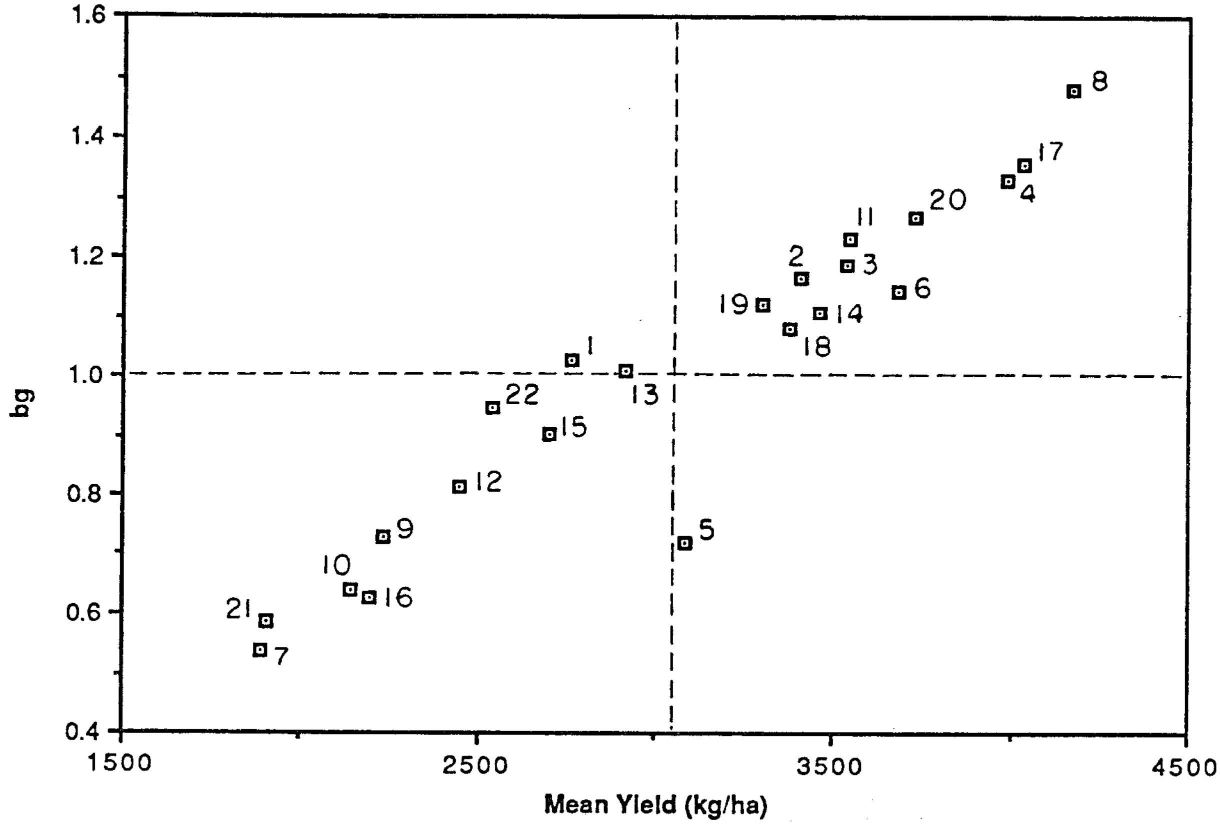

Figure 15. Plot of stability parameters against mean yield, Gujarat

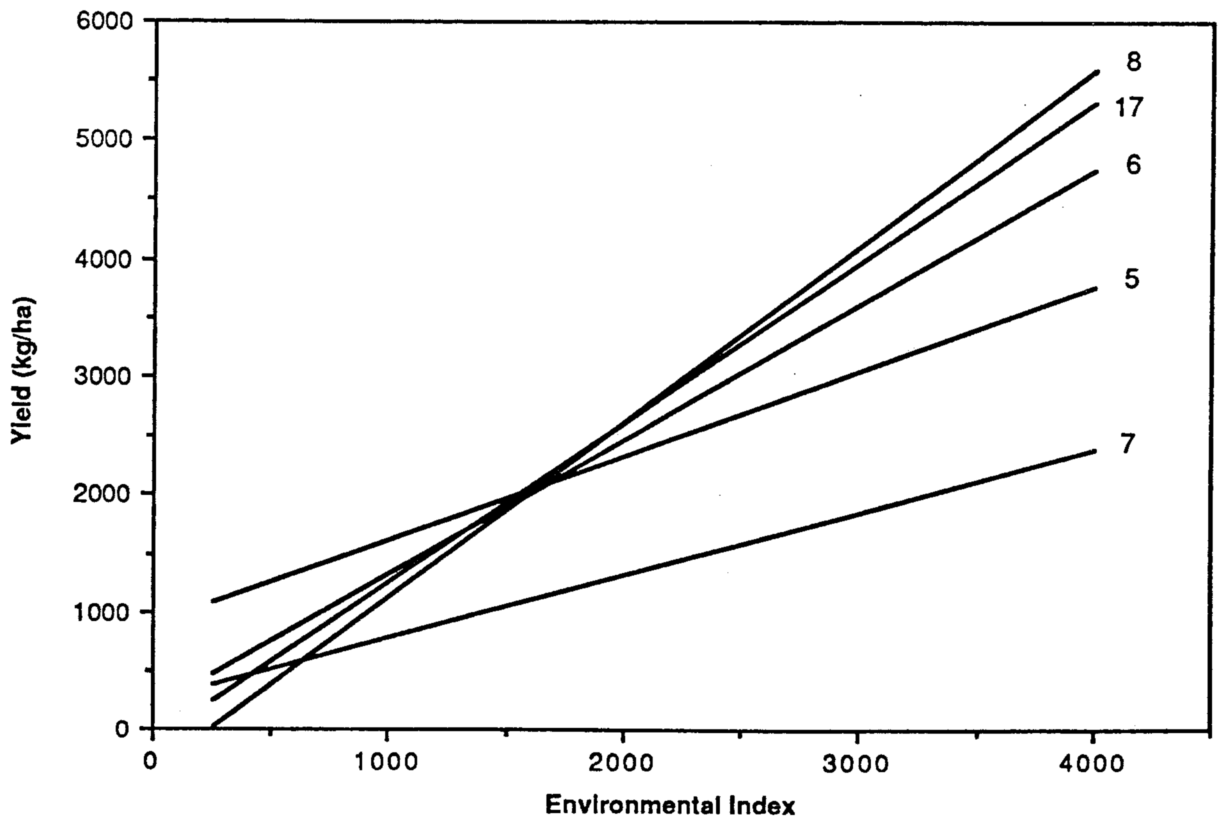

Figure 16. Genotype-environment interaction for selected genotypes, Gujarat

Evaluation Using Alternative Risk Decision Criteria

Based on this stability analysis, it is clear that the genotype recommendations are affected by the incorporation of independent meteorological information from the three sites. As suggested in sections *II* and *III*, the stability analysis does not provide a formal decision rule that either recognizes the probability of drought or deals with the trade-off between mean yield and stability. Thus, it is important to compare these results with those from the various decision criteria discussed in section *III*. These criteria are applied to the genotypes at each location and to the experiment using the predicted yields from the response function.

Entire efficiency frontiers are generated by mean-variance (EV), mean-Gini (MG) and first, second and third degree stochastic dominance (FSD, SSD and TSD) criteria (see section *III* for a discussion of the criteria).[6] Other methods reported in the table reduce the size of the efficient sets either by limiting the efficiency analysis to specified intervals of risk aversion (stochastic dominance with respect to a function, SDWRF)[7] or through a complete ranking according to certainty equivalents using the exponential utility, empirical moment generating function, EUMGF, from equation (19). The results from applying these criteria are summarized in Table 13. Selected data in support of the analyses are in Appendix D. Complete results are in Bailey (1988).

In developing the results in Table 13, it is important to recall that for the efficient frontiers associated with the first four criteria, no explicit values of the Arrow-Pratt risk aversion parameter (equation (4)) needed to be assumed. This was not the case, however, for either the SDWRF or the EUMGF criterion. That is, SDWRF can be used to order genotypes only if upper and lower bounds on the absolute risk aversion coefficient are specified. These bounds were derived from results reported by Binswanger (1978; 1980) of games where Indian farmers were asked to choose between a number of gambles each with two payoffs of equal probability. The derivation is in Appendix E.

In looking at the results in Table 13, several conclusions can be drawn. First, within each location, the results are more or less consistent across the alternative decision criteria. However, as was found in the stability analysis above, the sets of efficient genotypes differ

---

6 For each genotype, each observation on yield (each treatment in the trial, each year in the sample locations) is regarded as a single element of a discrete sample probability density function, each having equal probability of occurring. Anderson, Dillon and Hardaker (1977) report FSD, SSD and TSD rules for the case of a discrete probability distribution. They are comparable to the rules for continuous probability functions discussed above, except that the calculations involve summing over the cumulative functions rather than integration. Defining $\Delta x_i = x_i - x_{i-1}$ and if x_n is the highest value for x, then the discrete analogs of F_2 and F_3 are

$$F_2(x_r) = \sum_{i=2}^{r} F_1(x_{i-1})\,\Delta x_i \quad (r=2,\ldots,n)$$

$$F_2(x_1) = 0$$

and

$$F_3(x_r) = 1/2 \sum_{i=2}^{r} [F_2(x_i) + F_2(x_{i-1})]\Delta x_i \quad (r=2,\ldots,n)$$

$$F_3(x_1) = 0.$$

In this study, the analysis of first, second and third degree stochastic dominance was carried out using the Fortran program presented in Anderson, Dillon and Hardaker (1977, p. 313).

7 The SDWRF analysis was carried out using a modified form of a program written by Meyer, reported in King and Robinson (1981b) and modified by Loren Tauer.

Table 13. Efficient Genotypes for Alternative Approaches to Risk Analysis[a]

Criteria	Location			
	Experiment	Hyderabad	Anantapur	Gujarat
Stability	4, 18, 9, 16	8, 4, 5, 7	17, 6, 5, 16	8, 17, 4, 5, 7
EV	4, 18, 9, 16	8, 4, 5, 7	17, 6, 5, 16	8, 17, 4, 5, 7
Stochastic Dominance:				
FSD	4, 18, 14, 17, 6, 11, 13, 5, 9, 10, 7	8, 4, 17, 6, 20 18, 5	17, 6, 8, 3, 5, 4, 14, 18, 9, 15, 11, 16	8, 17, 4, 20, 6, 3, 14, 2, 18, 19, 5
SSD	4, 18, 6, 11, 5, 9	8, 4, 17, 6, 5	17, 6, 8, 3, 5, 4, 14, 9, 11	8, 17, 4, 6, 14, 18, 5
TSD	4, 18, 6, 11, 5, 9	8, 4, 17, 6, 5	17, 6, 5, 9	8, 17, 4, 6, 14, 18, 5
Mean-Gini (MG)[b]	4	8, 4	17, 6, 5	8
SDWRF:				
$r_A \leq 0$	4	8	17	8
$0 \leq r_A \leq .00005$	4	8	17	8
$.00005 \leq r_A \leq .00015$	4	8,4	17,6	8
$.00015 \leq r_A \leq .00037$	4	4	6	8, 17
$.00037 \leq r_A \leq .00198$	4, 18	4, 6	6, 5	17, 4, 6, 5
$.00198 \leq r_A \leq .005$	4, 18, 5, 9	6, 5	5, 9	4, 5
EUMGF (top 5 ranked, by CE)[b]				
$r_A = 0$	4, 18, 14, 17, 6	8, 4, 17, 6, 20	17, 6, 8, 3, 5	8, 17, 4, 20, 6
$r_A = .00005$	4, 18, 14, 17, 6	8, 4, 17, 6, 20	17, 6, 5, 8, 3	8, 17, 4, 20, 6
$r_A = .00015$	4, 18, 14, 17, 6	8, 4, 17, 6, 20	6, 17, 5, 3, 4	8, 17, 4, 6, 20
$r_A = .00037$	4, 18, 14, 6, 17	4, 17, 8, 6, 20	6, 5, 17, 4, 3	17, 4, 6, 8, 20
$r_A = .00198$	18, 9, 4, 14, 17	6, 5, 4, 18, 17	5, 9, 6, 18, 15	5, 6, 18, 4, 14
$r_A = .005$	9, 5, 18, 11, 16	5, 6, 8, 20, 17	9, 5, 6, 18, 16	4, 5, 6, 3, 18

[a]Numbers refer to genotypes. See Appendix A for descriptions. Genotypes are ranked according to their mean yield, except for EUMGF results which are ranked by certainty equivalents (CE's) as defined in equation (19).

[b]The data on which these criteria sets are determined are in Appendix D.

substantially across locations. This result supports the hypothesis that genotype evaluations must consider differences in specific local environmental conditions.

It is also interesting, although not too surprising, that the efficient genotypes as identified by the two simplest and most widely used risk criteria (the stability analysis and EV analysis) are identical for a given location. They all contain the lowest yielding, most stable genotypes. This consistency is undoubtedly explained in part by the fact that both criteria are based exclusively on mean yield and some measure of variation in yield. This is particularly true for the experiment and for Gujarat when the simple correlations between the standard deviations (Table D.1) and the regression slope coefficients from the stability analysis are 0.99 and 0.91. For the other two locations, these simple correlations are much lower.

Within each location, the results are also more or less consistent across the alternative approaches to risk analysis (Table 13). As expected, the efficient sets generated under FSD are quite large, containing more than 10 genotypes in the experiment and in two of the three locations. Only in Hyderabad are there fewer efficient genotypes. SSD analysis is effective in considerably reducing the set of genotypes under consideration to five genotypes in Hyderabad and seven in Gujarat, both of which have relatively less variable environments than Anantapur, which has a greater frequency of severe droughts and a SSD efficient set of nine genotypes. TSD is ineffective in reducing the size of the SSD set except in Anantapur. The inclusion of four more observations on yield in Anantapur did not affect the results from EV and MG analysis. However, in the stochastic dominance analysis, one more genotype (GNO2) is included in the FSD efficient set, and two more (GNO2 and GNO18) in the SSD efficient sets. The TSD efficient set remains unchanged.

The SSD efficient sets contain a larger number of the most highly ranked (in terms of mean yield) genotypes than the EV efficient sets, particularly in Hyderabad and Gujarat. This supports Anderson's (1974) proposition that as the environmental range becomes more limited, the greater the chance that only very highly ranked varieties will appear in the SSD or TSD sets.

The MG criterion exhibits far greater discriminatory power than the SSD and EV criteria. The MG criterion is a necessary condition for SSD. If distributions cross only once, then it is a sufficient condition. Examination of the simulated yield data revealed that many distributions cross more than once. The MG efficient set in Gujarat contains only one, the highest yielding, genotype. In Hyderabad the two highest yielding genotypes are MG efficient, while in Anantapur, the MG efficient set contains two high yielding genotypes plus a lower yielding, more stable genotype. It is not known, however, what the implications of this greater discriminatory power are, in terms of the degree of risk aversion. If the MG criterion is reducing the size of the SSD efficient set, it is doing so by rejecting some genotypes that may be preferred by some risk averse decision makers. The results correspond with those from the SDWRF analysis over the slight to moderate range of risk aversion and support Buccola's and Subaei's (1974) hypothesis that the MG criterion best represents the preferences of relatively weakly risk averse decision makers.

To obtain additional insight into this issue, similar comparisons can also be made between extended mean-Gini (MEG) efficient genotypes and those efficient by SDWRF.[8] This is possible because the absolute Gini index of equality $\delta(v) = \mu_F - \Gamma_F(v)$ is a weighted integration of the area under the CDF, $F(x)$, and a risk averse decision maker places more weight in the lower tail of the distribution by increasing the value of v.

Extended mean-Gini calculations were made by Bailey (1988) for integer values of v from 3 to 10, and for arbitrarily selected higher values of v = 20, 25 and 50. Selected results are presented in Table 14 and the tables in Appendix D (Tables D.7-D.10).[9]

Table 14. Extended Gini-efficient Genotypes[a]

V[b]	Experiment	Hyderabad	Anantapur	Gujarat
2	4	8, 4	17, 6, 5	8
3	4, 18	8, 4	17, 6, 5	8, 17
4	4, 18	8, 4	17, 6, 5	8, 17, 5
5	4, 18	8, 4	17, 6, 5	8, 17, 6, 5
6	4, 18	8, 4	17, 6, 5	8, 17, 6, 5
7	4, 18, 9	8, 4	17, 6, 5	8, 17, 6, 5
8	4, 18, 9	8, 4	17, 6, 5, 9	8, 17, 4, 6, 5
9	4, 18, 9	8, 4	17, 6, 5, 9	8, 17, 4, 18, 5
10	4, 18, 9	8, 4	17, 6, 5, 9	8, 17, 4, 18, 5
20	4, 18, 9	8, 4, 6, 5	n.a.[c]	8, 17, 4
25	4, 18, 9	8, 4, 6, 5	n.a.	n.a.
50	4, 18, 5, 9	8, 4, 6, 5	n.a.	n.a.

Source: Bailey 1988.

[a]Numbers in the table are genotypes as described in Appendix A.

[b]Exponent on the absolute parametric index from equation (10).

[c]Not applicable.

---

[8] Recall from section III that the necessary conditions for the distribution F to dominate G according to expected utility, where $\mu'(x) > 0$ and $\mu''(x) < 0$ are:

$$\mu_F \geq \mu_G \text{ and } \mu_F - \Gamma_F(v) \geq \mu_G - \Gamma_G(v) \text{ for } v \geq 1$$

where $\Gamma_F(v) = -v \operatorname{cov}[x,(1-F(x))^{v-1}]$, which is the MG criterion if v=2. For the extended mean-Gini, Shalit and Yitzhaki (1987) note that in the case of discrete distributions with K observations, one should calculate (MEG) by:

$$F(x_i) = [\text{Rank}(x_i) - 0.5]/K, \text{ and}$$

$$Z_i = [(K + .05 - \text{Rank}(x_i))/K]^{v-1} \text{ is the estimator of } [1-F(x_i)]^{v-1}.$$

[9] Some properties of the extended Gini should be noted: the extended Gini, $\Gamma(v)$, is non-decreasing in v, i.e., $d\Gamma(v)/dv \geq 0$; it follows that $\mu - \Gamma(v)$ is non-increasing in v. As v approaches infinity, $\mu - \Gamma(v)$ approaches a, the minimum outcome value (Shalit and Yitzhaki, 1984). The upper bound on v at which $\mu - \Gamma(v)$ approaches the minimum value, a from equation (10), varies over the four locations. At values above this upper bound the extended Gini misbehaves and is no longer non-decreasing in v.

From these results, it is clear that while the MG efficient set (v=2) from the trial corresponds to the risk intervals, $0 \leq r_A(x) \leq 0.00037$ in the SDWRF analysis, the mean-extended Gini (MEG) sets for v = 3, 4, 5, 6 correspond to the set, $0.00037 \leq r_A \leq 0.00198$. At the extreme value of v = 50, GNO5 is included and corresponds to interval, $0.00198 \leq R_A(x) \leq 0.005$. The union of all MEG efficient sets for v = 3, 4, ..., 50 corresponds to that from SDWRF for the full range of r_A. For Hyderabad, efficient sets from MG analysis and MEG analysis for values of v = 3, 4, ..., 10, contain GNO4 and GNO8 corresponding to SDWRF results for interval R2.[10] Only when $v > 10$, does the MEG set include GNO6 and GNO5, representing the higher risk averse intervals of R4 and R5.

In Anantapur, from the SDWRF analysis, GNO17 alone dominates over mild to moderate ranges of absolute risk aversion; from MG and MEG analysis, all values of $v \geq 2$ appear to represent severe or extreme risk aversion. In Gujarat the MG efficient set corresponds to those in intervals R0 to R2; a value of v = 3 appears to correspond to interval R3, and values of v = 5, 6 corresponds to intervals R4 and R5. At higher values of v, v = 9, GNO18 enters the MEG efficient set; under SDWRF analysis, GNO18 is always dominated by another genotype. The end result is that a single value of v cannot be used to evaluate genotypes for a given level of risk aversion in all locations; different values of v correspond to different values of $r_A(x)$ in the different locations and there is no way of specifying, *ex ante*, the degree of risk aversion represented by a particular value of v.

Somewhat in contrast to these efficiency criteria, the final alternative discussed is the EUMGF criterion. This alternative represents an attempt to circumvent the major criticism of stochastic dominance or any other efficiency criterion in which genotypes may be inefficient, but inefficient by only a small degree. The EUMGF approach is attractive in that by ranking genotypes in terms of their certainty equivalents, (equation (19)), it allows one to assess the extent of the differences in performance of genotypes at different levels of risk aversion. Only the top five ranked genotypes are presented for each specified level of r_A in Table 13. (Appendix Table D.6 contains the corresponding certainty equivalents to provide some idea of how close the rankings are; the full range of results is in Bailey (1988)). The results generally support those from SDWRF but also indicate the order in which genotypes enter and leave the efficient set as the degree of risk aversion increases. The inclusion of four more observations on yield in Anantapur does not affect the ordering of genotypes by SDWRF except at extreme levels of risk aversion ($0.00198 < r < 0.005$), when GNO6 remains in the SDWRF efficient set. Similarly, in the EUMGF analysis, the positions of GNO5 and GNO6 are inverted when $r = 0.005$.

To interpret the results from SDWRF and EUMGF analysis, however, it is important to remember that the values for r_A were generated from an experiment by Binswanger. He finds that when payoffs in the game played with farmers rise to levels equivalent to agricultural investments (as in the 50 RS game from which the bounds on r_A used in this analysis are derived), most farmers have similar pure attitudes toward risk and were largely concentrated in the moderate and intermediate risk aversion classes (35% and 40%, respectively), with approximately 10% in the risk neutral and risk preferring classes and 8% in the severe and extreme classes (Binswanger 1978, p. A-55).

One could, therefore, argue that if any selection of genotypes is to be made, it should be based on preferences in the intermediate and moderate risk aversion classes. If our decision were based on the SDWRF results for the trial only, then GNO4 would be identified as the preferable genotype. If the location-specific results from the sample sites are used, then GNO4 would be selected in Hyderabad, GNO6 in Anantapur, and GNO8 in

---

[10] RI refers to risk interval I defined in Appendix E. As I increases, the interval reflects greater aversion to risk.

Gujarat; the more stable, but lower yielding, genotypes such as GNO5 and GNO9 are excluded from consideration when attention is focused on the moderate and intermediate levels of risk aversion. The same conclusions would be drawn from the results from EUMGF analysis.

Binswanger's intervals of risk aversion were, however, determined by the game, and farmers were classed accordingly; the bounds on r_A were not elicited directly from farmers. The rescaling of the values of r_A for use in our analysis involved generalizations regarding average area, price, *etc.* We have also, inherently, assumed that farmers attitudes towards the payoffs in the game are the same as towards actual returns from a groundnut crop. In short, the bounds on r_A used in the analysis must be regarded as more or less arbitrary. An alternative use of the SDWRF and EUMGF approaches would be to conduct a search over the intervals of r_A, sequentially sub-dividing each interval, to pinpoint the exact values of r at which the ordering of genotypes changes. However, it is unlikely that researchers will ever be able to measure either individuals' risk preferences or the distribution functions of outcomes with the accuracy required for such a search. Thus, it is argued that the results contained here are sufficient to indicate how the ordering of genotypes changes as risk aversion increases.

Furthermore, care must be used in the EUMGF approach because it is based on the assumption of constant absolute risk aversion associated with the negative exponential utility function. Binswanger (1980, p. 400) finds evidence that suggests farmers have non-linear, risk averse utility functions which exhibit increasing partial risk aversion. Zeckhauser and Keeler (1970) show that constant absolute risk aversion utility functions display partial risk aversion; increasing partial risk aversion is satisfied by the negative exponential utility function $U(w) = -e^{-kw}$, where $w = W+x$, the argument of the partial risk aversion function. Therefore, the use of the EUMGF approach to order genotypes according to risk preferences of farmers such as those included in Binswanger's study is not inappropriate.

VIII - SUMMARY AND CONCLUSIONS

This study is based on the premise that a major source of risk facing farmers in the semi-arid tropics is the year-to-year variation in crop production. Uncertain yields are due to the variable environment and particularly to fluctuations in available moisture, both within and between years. Any evaluation of the performance of genotypes requires observations on yields over a range of environments. The traditional approach has been to use the results from multisite, multiseason (MSMS) nursery trials, and to regress individual genotype yields, at each site in each season, on an environmental index composed of mean genotype yields.

Farmers, however, are concerned with how new genotypes will perform in their areas. The presumption that results from MSMS trials are good proxies for performance over time in any given location is questionable. Measurements of environmental factors at each trial site are rarely recorded and the full range of possible environmental conditions may not be covered by the trial. Results from sites in which the crop has failed also may be disregarded or go unreported. Furthermore, the environmental index reflects the totality of environmental factors as represented by mean genotype yields; the environments (sites) are, in effect, classified according to average genotype performance in that environment. Results cannot, therefore, be extrapolated to other locations as no independent measure of the environment exists. Finally, while the approach classifies genotypes by their mean yield and relative stability, without further information on farmers' preferences, a choice of genotype according to yield and degree of stability can not be made. Any method of evaluating a new technology, in terms of its degree of riskiness, must take account not only of the probability distribution of returns, but also of farmers' preferences and attitudes toward risk.

To address some of these shortcomings in this traditional approach, the objectives of this study are: a) to explore alternative ways of generating data on genotype performance which include independent information on environmental factors specific to production regions, and b) to compare the plant scientists' "stability" analysis for different approaches to evaluating genotype performance with alternative risk decision criteria that account for both the probability of adverse environmental conditions, as well as different attitudes toward risk. An application is made to selecting risk reducing groundnut genotypes in India. Emphasis is placed on evaluating genotypes in terms of their response to water stress because the performance of rainfed crops in the semi-arid tropics is largely determined by moisture availability.

To accomplish these objectives, the results from a single site experiment conducted by ICRISAT, in which 22 genotypes of groundnut were subjected to a range of drought conditions, are used to estimate the relationships between yield and available moisture. Groundnut genotypes of comparable maturity were selected to include lines found to be tolerant, average or susceptible to drought in previous screenings. These included established commercial and Indian cultivars, as well as advanced breeding lines. The drought patterns varied in the timing, duration and intensity of water stress, ranging from non-stressed control plots to plots in which the crop received virtually no water for the duration of the growing season.

In selecting the study locations, attention was focused on the two major groundnut-producing regions in India: Gujarat and Andhra Pradesh. Between them, these two states accounted for just under half of India's groundnut production in 1984. Based on the availability of extensive historical meteorological data, one site was selected from Gujarat and two sites, Hyderabad and Anantapur, from Andhra Pradesh. The sites differ climatically and for the latter two sites, about 80 years of daily meteorological data were available, while for Gujarat about 30 years were available.

Simulating Yield Response

To affect control of the drought treatments through the application of irrigation without interference from rainfall, the trial was planted in the late post rainy season and continued into the summer season under conditions that, meteorologically, are different to those in the rainy season. Furthermore, the trial is characterized by discrete applications of water, irrigation being applied only when symptoms of wilting were observed in the non-stressed control plots. This is in contrast to the nature of rainfall in the rainy season, when soil moisture may be continually replenished by intermittent rain showers. Thus, the problem of simulating yields in each of the three sample locations is compounded by the difficulties involved in translating results from the post rainy season trial to rainy season conditions. To do this, an index of relative water availability (RAW, available moisture relative to potential water requirements, $0 \leq RAW_i \leq 1$) was developed using a simple daily soil moisture budgeting procedure. It takes account of the soil moisture holding capacity and the consumptive water use requirements of the plant, given the potential evapotranspiration rates and stage of crop growth, and allows for the carryover of soil moisture from one period to another. To account specifically for the effect of timing of water application, the growing season was divided into four growth phases: seedling-flowering; pegging to pod set; pod set to pod filling; pod filling to maturity. The relationships between yield and relative available water (RAW_i) in the four phases ($i=1...4$) were estimated using data from the experiment. The estimated coefficients from each genotype specific response function are used, along with estimates of RAW from the meteorological data, to simulate yields for a series of years in each location.

A number of difficulties were encountered in modelling the response of groundnut to RAW_i in different growth stages. Modelling the response to drought in the early vegetative and flowering phase was particularly problematic. Groundnut has the ability at this stage of growth to lie dormant until soil water is replenished. Initial analysis of the experimental data indicated that, over some range, water stress in the early phase actually has a positive effect on yield, *i.e.*, marginal productivity of RAW in this early phase may be negative. Moreover, sensitivity to droughts later in the season is modified according to whether or not drought occurred in the early phase. Further, drought treatments in the experiment did not cover the full range of $0 \leq RAW_i \leq 1$ in all growth phases ($i=1...4$) and difficulties were encountered with predicting outside the range of drought conditions in the trial.

Initially a number of possible response functions were estimated, but because of the confounding effect of stress in the early phases of plant development, the translog flexible form was the only one that would account for the interaction between relative water availability at different stages of growth and prohibits estimated yields from being negative.

The difficulties encountered in generating yield distributions lead to some suggestions regarding future experimentation and research. The problems involved in translating results from the post rainy season to the rainy season need to be resolved. Ideally, conducting similar trials in the rainy season, if a satisfactory means of creating variable stress treatments can be found, would resolve many of the problems. Alternatively, research efforts could be directed more toward agrometeorological investigations. Further development of crop-weather models would allow the characterization of groundnut growing environments. As was discussed earlier, in relation to the traditional approach to stability analysis, environments cannot be classified according to the environmental index used in that approach, as the index is determined by average crop performance. Similarly, in the experiment, the treatments (timing of water application) were, in effect, determined by the crop, water being applied only at the first signs of wilting in the non-stressed control plots. If, however, a more satisfactory model of the factors determining relative available water in any given environment can be developed, then it can be ensured that treatments in the trial adequately represent the full range of conditions prevailing in the sample locations, and the problem of predicting outside the range would be resolved. Classification of groundnut growing environments in terms of independent meteorological and other environmental factors would assist in the extrapolation of results from a suitably designed single site trial in which the same factors had been measured.

Risk Analysis and the Evaluation of Genotype Performance

Having generated yield distributions for three sample locations, the next task was to examine a variety of approaches to evaluating genotypes in terms of their yield and degree of stability or, conversely, their riskiness. The traditional approach to "stability" analysis is compared to approaches involving various efficiency criteria such as EV analysis, first, second and third degree stochastic dominance (FSD, SSD and TSD) and stochastic dominance with respect to a function (SDWRF), and the more recently developed mean-Gini (MG) and mean-extended Gini (MEG) criteria, and to the exponential utility empirical moment generating function approach (EUMGF) that gives a complete ordering of alternative actions for specified levels of absolute risk aversion.

Within any given location, results are consistent across the range of methods used. Comparing "stability" analysis with EV analysis, the plots of the stability parameter, against mean yields, and the plots of standard deviations against mean yield, are strikingly similar. The empirical results indicate that the efficient set produced by EV and MG analysis are subsets of the SSD efficient set. Other methods reduce the size of the efficient set, either by restricting the efficiency analysis to specified intervals of risk aversion (SDWRF and MEG), or through a complete ranking (EUMGF). With such consistency of results, the choice of a particular approach to the analysis of risk efficiency depends on available computational facilities and on one's hypotheses regarding the risk attitudes of producers.

If we accept Binswanger's finding that most farmers are only moderately risk averse, then the MG criterion is perhaps appropriate, bearing in mind that it appears to best represent weak risk aversion. It is simple to apply and produces a much smaller set than the EV or SSD criteria. The results coincide with those from SDWRF over the slight to moderate, or intermediate, risk aversion intervals. If, however, the objective is to identify genotypes to be released to a more heterogenous group of farmers, including a range of moderately to more severely risk averse individuals, as represented in the SDWRF analysis, then the problems associated with the interpretation of the risk parameter in the MG and MEG analyses detract from their use. Under such conditions, SDWRF reduces the size of the SSD and TSD efficient sets at most locations.

One criticism of efficiency criteria such as SSD or SDWRF concerns the accuracy of their discriminatory power; it may be that genotypes that are identified as inefficient are inefficient by only a small degree. The EUMGF approach is attractive in that, by ranking genotypes in terms of their certainty equivalents, it allows the researcher to assess the extent of the differences in performance of genotypes at specified levels of risk aversion (if one accepts the underlying assumption of constant absolute risk aversion, or increasing partial risk aversion). For example, in the EUMGF analysis of the experiment yields, GNO14 ranks among the top four genotypes over neutral to severe risk aversion, yet it is not included in the SSD efficient set, and was identified as inefficient over all levels of risk aversion in the SDWRF analysis. The value of results from both the SDWRF and EUMGF approaches, however, depends on the specification of the absolute risk aversion coefficient.

If comparisons are made across sites, the trial results are very different from those for the three sample locations. Based on the trial results alone, one may speculate that high yielding genotypes such as GNO4 and GNO18 would be suitable for areas with a relatively low probability of water stress, while the greater stability of genotypes such as GNO5 and GNO9 would make them more attractive in areas more prone to drought. However, GNO18 does not appear in the efficient sets in any of the sample locations. From the application of other methods of risk analysis to the experiment yields, GNO4 would appear to be the dominant genotype even at higher levels of risk aversion; from the analysis of the trial results alone, there is no indication of the prominence of GNO8 found in the analysis of simulated yields in Hyderabad and Gujarat.

In the final analysis, how does the inclusion of independent meteorological information and economic concepts of risk compare with results from stability analysis of the single-site trial results alone? First, the results show that the incorporation of independent information allowing the simulation of yield distributions leads to different conclusions than would have been drawn from the trial results alone. Second, the incorporation of different assumptions regarding risk leads to different orderings of genotypes. While high yielding genotypes are preferred over the moderately risk averse range, they are not dominant over the full range of risk aversion.

The analysis and results presented in section *VII* depend crucially on the simulation of yields in each location and, therefore, on the estimation of the response of yields to relative water stress. Given the consistency of results within each location, it is suggested that the direction of most concern for future research lies not with the choice of selection criteria, but with the improved measurement and modelling of agrometeorological relationships.

Risk Efficient Genotypes for Three Sites in India

A specific objective of this study is to assess how a group of groundnut genotypes would perform in three sample locations in India. The analysis provides insights into the relative performance of the various genotypes, but any decisions regarding selection of specific genotypes are left to the ICRISAT plant breeders who are far more familiar with the characteristics of each genotype. However, a consideration of the results, for varying degrees of risk aversion, for the three sample locations allows us to offer some general recommendations for the selection of genotypes for further screening. The set of risk-efficient genotypes which warrant further attention in each location are as follows:

Hyderabad: GNO4, GNO5, GNO6, GNO8, GNO17;

Anantapur: GNO5, GNO6, GNO9, GNO17;

Gujarat: GNO4, GNO5, GNO6, GNO8, GNO17.

Any further reductions of the efficient sets are hampered by the need to specify farmers attitudes toward risk.

GNO8 appears to be particularly responsive to favorable conditions, while GNO5 and GNO9 perform better in lower yielding environments. A number of genotypes appear in the risk efficient sets for all locations and may be regarded as widely adaptable, as well as risk efficient, for specific intervals of risk aversion, in any given location.

GNO4 (ICGS-36), GNO5 (ICGS-11), and GNO6 (ICGS-35) are all ICRISAT breeding lines. GNO5 was recommended, in 1985, for release for post rainy season cultivation in central India and the peninsula, and was released (as Robut 33-1-18-8-B1) in 1986 (ICRISAT, 1986, p.367). GNO8 (X41-x-1-B x Goldin-1) and GNO9 (Manfredi x X-14-4-B-19-B) are both crosses between established commercial cultivars. GNO17 (EC 109271 (55-437)), of West African origin, has already been identified as drought tolerant in a variety of trials.

The analysis provides some surprising results with regard to the relative performance of such well established genotypes as TMV-2 (GNO10), J11 (GNO12), and J24 (GNO22). These are genotypes that have already been released to the farming community, have been widely adopted and are frequently used as control genotypes in ICRISAT trials. They are among the lowest ranked genotypes (in terms of mean yield), particularly in Anantapur; none of them appear in the SSD efficient sets in any of the sample locations.

Concluding Observations

This study evaluates genotypes with respect to their response in pod yield to variations in timing, duration and intensity of water stress. In a more comprehensive evaluation, however, other properties of the genotypes should also be considered. The set of genotypes tested vary considerably in their degree of susceptibility to various diseases and pests. No account is taken here of other variables such as soil characteristics, or other inputs such as fertilizer, and disease and pest control. Yield response to variations in uncontrollable factors such as available moisture may well be affected by variations in input levels. The treatments in the trial received uniform applications of fertilizer and herbicide; production costs were constant across all treatments. Consequently, yields were used as a proxy for gross returns. The inclusion of other variables would allow for variation in production costs to be incorporated in the analysis and genotypes could then be evaluated in terms of gross margins.

Farmers employ a range of management strategies to reduce the risk of income loss due to unpredictable variations in yield. One approach to spreading risk is through diversification, either varietal diversification within a crop, crop diversification, or through intercropping. Recommendation of a number, or mix, of genotypes and the encouragement of diversification of plant stands may aid in the reduction of production variability due to climatic variability or pests and diseases. In this study, the selection of genotypes is treated as a choice between mutually exclusive alternatives. No attempt is made in this analysis to construct efficient "portfolios" of genotypes. Such an extension would be relatively simple using results from the EV, MG, or MEG analysis. It is not so easily achieved using SSD or SDWRF criteria, where pairwise comparisons must be made between all proposed combinations of genotypes.

Constructing portfolios would be most important in evaluating trade-offs between drought escape and drought resistance. However, these would almost always involve genotypes of different duration because shorter duration material has a greater chance of escaping late season droughts while longer duration genotypes are usually more drought tolerant and have higher potential yields. Intercropping genotypes of different maturity could be one approach to achieving yield gains in locations such as Gujarat where end of season drought is frequent. Research on this question is currently being conducted by the groundnut scientists at ICRISAT.

References

Anderson, J. R. "Risk Efficiency in the Interpretation of Agricultural Production Research". Review of Marketing and Agricultural Economics 42(1974):131-184.

Anderson, J. R., J. L. Dillon, and B. Hardaker. Agricultural Decision Analysis. Ames, Iowa: Iowa State University Press, 1977.

Arnon, I. "Physiological Principles of Dryland Crop Production". In U. S. Gupta, (ed) Physiological Aspects of Dryland Farming, New Delhi: Oxford & IBH Publishing Co, 1975.

Arrow, K. J. Aspects of the Theory of Risk Bearing. Helsinki: Yrjo Jahnssonin Saatio, 1965.

Baier, W. "Crop-Weather Models and Their Use in Yield Assessment". WMO Technical Note No. 151, WMO No. 458. World Meteorological Organization, Geneva, 1977.

Baier, W. and G. W. Robertson. "A New Versatile Soil Moisture Budget". Canadian Journal of Plant Science. 46(1966):299-315.

Bailey, E. The Use of Risk Analysis in the Evaluation of Genotype Performance in Drought Prone Areas. Unpublished Ph.D., Cornell University, 1988.

Barah, B. C., H. P. Binswanger, B. S. Rana, and N. G. P. Rao. "The Use of Risk Aversion in Plant Breeding; Concept and Application". Euphytica 30(1981): 451-458.

Baumol, W. J. "An Expected Gain-Confidence Limit Criterion for Portfolio Selection". Management Science. Vol. 10, no. 1, 1963.

Bey, R. P. and K. M. Howe. "Gini's Mean Difference and Portfolio Selection: An Empirical Evaluation". Journal of Financial and Quantitative Analysis. 19(1984):329-338.

Binswanger, H. P. "Risk Attitudes of Rural Households in Semi-Arid Tropical India". Economic and Political Weekly. Review of Agriculture, June 1978: A49-A62.

______________. "Attitudes Toward Risk: Experimental Measurement in Rural India". American Journal of Agricultural Economics. 62(1980):395-407.

Binswanger, H. P. and B. C. Barah. "Yield Risk, Risk Aversion, and Genotype Selection: Conceptual Issues and Approaches" ICRISAT Research Bulletin No. 3, International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India, 1980.

Boisvert. R. N. "Choice in Risk Situations: A Comparison and Evaluation of Alternative Criteria". Staff Paper No. 72-14, Dept. of Agricultural Economics, Cornell University, Ithaca, New York, 1972.

______________. "The Translog Production Function: Its Properties, Its Several Interpretations and Estimation Problems". A. E. Res 82-28, Dept. of Agricultural Economics, Cornell University, Ithaca, New York, 1982.

______________. "The Role of Alternative Risk Programming Models in Empirical Research". Staff Paper, No. 85-5, Dept. of Agricultural Economics, Cornell University, Ithaca, New York, 1985.

Boote, K. J., J. R. Stansell, A. M. Schubert, and J. F. Stone. "Irrigation, Water Use, and Water Relations". In H. E. Pattee and C. T. Young (eds) Peanut Science and Technology. American Peanut Research and Education Society, Inc., Yoakum, Texas, 1982.

Buccola, S. T. and A. Subaei. "Mean-Gini Analysis, Stochastic Efficiency and Weak Risk Aversion". The Australian Journal of Agricultural Economics 28(1984):77-86.

Campbell, G. S. and R. Diaz. "Simplified Soil-Water Balance Models to Predict Crop Transpiration". Paper presented at the International Consultants' Meeting on Research on Drought Problems in the Arid and Semi-Arid Tropics, November 17-20, 1986. International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India, 1986.

Central Statistical Organization. "Statistical Abstract, India", Dept. of Statistics, Ministry of Planning, Govt. of India, New Delhi, 1984.

Cochran, M., W. Lodwick, and L. Robison. "Rankings of Action Choices Under Uncertainty: A Discussion of Stochastic Dominance Techniques". Staff Report 1982-68, Dept. of Agricultural Economics, Michigan State University, East Lansing, 1982.

Collender, R. N. and D. Zilberman. "Land Allocation Under Uncertainty for Alternative Specifications of Return Distributions". American Journal of Agricultural Economics 67(1984):779-786.

Collender, R. N. and J. A. Chalfant. "An Alternative Approach to Decisions Under Uncertainty Using the Empirical Moment-Generating Function". American Journal of Agricultural Economics 68(1986):727-731.

Day, R. H. "Probability Distributions of Field Crop Yields". Journal of Farm Economics 47(1965):713-741.

Directorate of Economics and Statistics. "Area and Production of Principle Crops in India, 1981-1984". Dept. of Agriculture and Cooperation, Ministry of Agriculture, Govt. of India, New Delhi, 1984.

Eberhart, S. A. and W. A. Russell. "Stability Parameters for Comparing Varieties". Crop Science 6(1966):36-40.

Evenson, R. E., J. C. O'Toole, R. W. Herdt, W. R. Coffman, and H. E. Kauffman. "Risk and Uncertainty as Factors in Crop Improvement Research". IRRI Research Paper Series, Number 15, International Rice Research Institute (IRRI), Los Banos, 1978.

Finlay, K. W. and G. N. Wilkinson. "The Analysis of Adaptation in a Plant-Breeding Programme". Australian Journal of Agricultural Research 14(1963):742-754.

Fishburn, P. C. and R. G. Vickson. "Theoretical Foundations of Stochastic Dominance". In Whitmore, G. A. and M. C. Findlay (eds) Stochastic Dominance. An Approach to Decision Making Under Risk. Lexington Books, USA, 1978.

Freeman, G. H. and J. M. Perkins. "Environmental and Genotype-Environmental Components of Variability. VIII. Relations Between Genotypes Grown in Different Environments and Measures of These Environments. Heredity 27(1971):15-23.

Hadar, J. and W. R. Russell. "Rules for Ordering Uncertain Prospects". American Economic Review 59(1969):25-34.

______________. 1971. "Stochastic Dominance and Diversification". Journal of Economic Theory 3(1971):288-305.

Hammond, J. S. "Simplification of Choice Under Uncertainty". Management Science, 20(1974):1047-1072.

Hanks, R. J., J. Keller, V. P. Rasmussen, and G. D. Wilson. "Line Source Sprinkler for Continuous Variable Irrigation". Crop Production Studies, Journal of the Soil Science of America, 40(1976):426-429.

Hanks, R. J. and V. P. Rasmussen. "Predicting Crop Production as Related to Plant Water Stress". Advances in Agronomy 35(1982):193-215.

Hanoch, G. and H. Levy. "The Efficiency Analysis of Choices Involving Risk." Review of Economic Studies 36(1969):335-346.

Hardwick, R. C. and J. T. Wood. "Regression Methods for Studying Genotype-Environment Interaction". Heredity 28(1972):209-222.

Hazell, P. B. R. "Overview of Issues. Introduction". In P. B. R. Hazell (ed) Summary Proceedings of a Workshop on Cereal Variability. International Food Policy Research Institute (IFPRI), Washington, D.C., 1986.

Henderson, J. M. and R. E. Quandt. Microeconomic Theory. Third Edition, New York: McGraw-Hill, 1980.

Hexem, R. W. and E. O. Heady. Water Production Functions for Irrigated Agriculture. Ames, Iowa: Iowa State University Press, 1978.

Jodha, N. S. "Resource Base as a Determinant of Cropping Patterns". Occasional Paper 14, Economics Department, International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India, 1977.

Kendall, M. G. and A. Stuart. "The Advanced Theory of Statistics". Vol. 1. Distribution Theory. London: Charles Griffin and Co. Ltd, 1958.

King, R. P. and L. J. Robison. "An Interval Approach to Measuring Decision Maker Preferences". American Journal of Agricultural Economics 63(1981a):510-520.

______________. "Implementation of the Interval Approach to the Measurement of Decision Maker Preferences". Michigan State University Agricultural Experiment Station Research Report 418, 1981b.

Knight, R. "The Measurement and Interpretation of Genotyoe-Environment Interactions". Euphytica 19(1970):225-235.

McGuckin, J. T., C. Mapel, R. Lansford and T. Sammis. "Optimal Control of Irrigation Scheduling Using a Random Time Framework". American Journal of Agricultural Economics 69 (1987) 123-133.

Menezes, C. F. and D. L. Hanson. "On the Theory of Risk Aversion". International Economic Review 11 (1970):481-487.

Meyer, J. "Increasing Risk". Journal of Economic Theory 11(1975):119-132.

______________. "Choice Among Distributions". Journal of Economic Theory 14(1977a): 326-336.

______________. "Second Degress Stochastic Dominance with Respect to a Function". International Economic Review 18 (1977b):477-487.

Minhas, B. S., K. S. Parikh, and T. V. Srinivasan. "Toward the Structure of a Production Function for Wheat Yields with Dated Inputs of Irrigation Water". Water Resources Research 9(1973):383-393.

Moll, R. H. and C. W. Stuber. "Quantitative Genetics - Empirical Results Relevant to Plant Breeding". Advances in Agronomy 26(1974):277-313.

Nageswara Rao, R. C., and J. H. Williams. "Effects of Single and Multiple Droughts with Varied Timing, Intensity and Duration on Groundnuts (Arachis hypogaea L.)." I. General Responses. International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India. (mimeo), 1985.

Nageswara Rao, R. C., J. H. Williams and Murari Singh. "Effects of Single and Multiple Droughts with Varied Timing, Intensity and Duration on Groundnuts (Arachis hypogaea L.)" II. Relation Between Pod Yield Potential and Sensitivity to Drought. International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India. (mimeo), 1985.

Ojomo, O. A. and B. O. Adelana. "Variety x Environment Interactions in Groundnut (Variety Tests in Western Nigeria)". Journal of Agricultural Science 75(1970):419-420.

Ong, C. K. "Agroclimatological Factors Affecting Phenology of Groundnut". In Agrometeorology of Groundnut. Proceedings of an International Symposium, Aug. 21-26, 1985, ICRISAT Sahelian Center, Niamey, Niger. International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India, 1986.

Pindyck, R. S. and D. C. Rubinfeld. Econometric Models and Economic Forecasts. Second Edition. New York: McGraw-Hill, 1981.

Pope, R. D. and R. F. Ziemer. "Stochastic Efficiency, Normality and Sampling Errors in Agricultural Risk Analysis". American Journal of Agricultural Economics 66 (1984):31-40.

Quandt, R. E. and J. B. Ramsey. "Estimating Mixtures of Normal Distributions and Switching Regressions". Journal of the American Statistical Association 73(1978): 730-741.

Raskin, R. and M. J. Cochran. "Interpretations and Transformations of Scale for the Pratt-Arrow Absolute Risk Aversion Coefficient: Implications for Generalized Stochastic Dominance. Western Journal of Agricultural Economics 11(1986):204-210.

Shalit, H. and S. Yitzhaki. "Mean-Gini, Portfolio Theory and the Pricing of Risky Assets". The Journal of Finance 35 (1984): 1449-1468.

______________. "Evaluating the Mean-Gini Approach to Portfolio Selection". (mimeo), 1987.

Stern, R. D. "Interpreting the Results from Irrigation Experiments". Paper presented at the International Consultants' Meeting on Research on Drought Problems in the Arid and Semi-Arid Tropics, November 17-20, 1986, International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India, 1986.

Whitmore, G. A. "Third-Degree Stochastic Dominance". American Economic Review 60(1970):457-459.

Williams, J. H., R. C. Nageswara Rao, R. Matthews and D. Harris. "Response of Groundnut Genotypes to Drought". In Agrometeorology of Groundnut. Proceedings of an International Symposium, Aug 21-26, 1985, ICRISAT Sahelian Center, Niamey, Niger. International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.WMO. 1975. Drought and agriculture. WMO Technical Note No. 138, WMO No. 392. World Meteorological Organization, Geneva, 1986.

Yassour, J., D. Zilberman, and G. C. Rausser. "Optimal Choices among Alternative Technologies with Stochastic Yield". American Journal of Agricultural Economics 63(1981):718-723.

Yates, F. and W. G. Cochran. "The Analysis of Groups of Experiments". Journal of Agricultural Science 28(1938): 556-580.

Yitzhaki, S. "Stochastic Dominance, Mean Variance, and Gini's Mean Difference". American Economic Review 72 (1982):178-185.

______________. "On an Extension of the Gini Inequality Index". International Economic Review 24 (1983):617-628.

Zeckhauser, R. and E. Keeler. "Another Type of Risk Aversion". Econometrica 38 (1970): 661-665.

APPENDICES A - E

Table A.1. Characteristics of Genotypes included in the Trial

GNO	ICG	GNP#	Identity	Variety	Comments
1		#35	CGC-4063	SP	Drought tolerant in screening trials in 1981-83
2		#402	J 11 x Robut 33-1	SP	High yielding cross between two established Indian cultivars
3		#404	ICGS-24	SP	ICRISAT advanced breeding line
4		#405	ICGS-36	SP	ICRISAT advanced breeding line; yielded > 6,000kg./ha. in 1981 trials
5		#406	ICGS-11	SP	ICRISAT advanced breeding line; included in kharif AICORPO CVT in Gujarat, 1984, 1985
6		#408	ICGS-35	SP	ICRISAT advanced breeding line; yielded > 6,000kg./ha. in 1981 trials
7		#409	ICGS-21	SP	ICRISAT advanced breeding line; yielded > 6,000kg./ha. in 1981 trials; included in kharif AICORPO CVT in Gujarat, 1984, 1985, promoted to rabi AICORPO NET, 1985
8		#411	X41-x-x-1-B x Goldin-1	SP	High yielding in 1982 trials
9		#413	Manfredi x X-14-4-B-19-B	SP	Cross between two established high yielding cultivars; X-14-4-B-19-B is used in breeding programs to increase nitrogen fixation
10	221	#87	TMV 2	SP	Long established Indian selection; susceptible to infection by *A. flavus*, and highly susceptible to foliar disease,
11	1104	#94	Faizapur 1-5-2	SP	High yielding Indian cultivar; moderately resistant to infection by *A. flavus*.

Table A.1. (cont.)

GNO	ICG	GNP#	Identity	Variety	Comments
12	1326	#75	J 11	SP	Long established high yielding Indian cultivar; susceptible to foliar disease, resistant to *A. flavus*.
13	1697	#42	NC Ac 17090	VAL	From Peru; drought tolerant breeding line, resistant to foliar disease, susceptible to *A. flavus*.
14	1712	#700	NC Ac 17142	VAL	From Brazil; extreme photo-period sensitivity; resistant to foliar disease
15	2738	#73	Gangapuri	VAL	Established Indian cultivar; susceptible to PMV, pod rot and infection by *A. flavus*.
16	3605	#221	EC 76444	VAL	Recent acquisition
17	3657	#38	EC 109271 (55-437)	VAL	Early maturing Senagalese genotype; drought tolerant in screening trials, 1981-83
18	3704	#214	EC 21024	VAL	Recent acquisition
19	3730	#129	Manfredi-107	SP	High yielding commercial cultivar
20	4790	#84	Krapovicas Str#16	VAL	From Argentina; identified as drought tolerant in trials, 1981-83, relatively resistant to foliar diseases
21	5094	#239	NC Ac 16129	VAL	Promising line from Brazil; performed well in ICRISAT yield trials, 1983
22	7827	#869	JL 24	SP	High yielding Indian cultivar; susceptible to foliar diseases

Source: Nageswara Rao, Williams and Singh (1985).

Notes: ICG - ICRISAT Groundnut Germplasm number; GNP = ICRISAT Groundnut Program number; GNO = number used to identify genotypes in this study; SP = Spanish (variety vulgaris); VAL = Valencia (variety fastigiata), NC Ac = North Carolina accession number; EC = original import number; AICORPO = All India Coordinated Research Project on Oilseeds; CVT = Coordinated Varietal Trial; NET = National Elite Trial

Table B.1 Marginal Products from Translog Response Functions, Estimated for Selected Values of RAW_i

GNO[a]	RAW_i = mean, i = 1,2,3,4				RAW_i = maximum value, i = 1,2,3,4			
	RAW_1	RAW_2	RAW_3	RAW_4	RAW_1	RAW_2	RAW_3	RAW_4
1	-114	198	286	132	-59	210	287	257
2	-83	230	375	115	52	252	416	222
3	-55	143	393	104	-8	161	504	170
4	-68	166	438	123	-37	195	584	194
5	-243	114	310	96	-437	166	387	89
6	-47	154	373	125	-22	227	514	184
7	-504	165	220	71	-535	102	94	79
8	-61	160	395	122	-21	268	733	252
9	-144	96	204	110	-179	112	198	148
10	-638	237	272	58	-774	174	146	86
11	-239	180	394	125	-323	243	522	215
12	-183	196	237	98	-200	223	220	139
13	-258	192	338	125	-368	153	295	180
14	-207	186	395	116	-175	184	446	185
15	-49	64	312	90	29	57	386	150
16	-23	67	232	67	74	64	259	103
17	-58	223	392	117	-41	309	571	238
18	-151	118	370	134	-158	126	438	200
19	-62	179	356	77	-50	230	484	151
20	-17	155	387	129	156	177	575	242
21	-323	138	206	99	-304	71	83	133
22	112	177	249	92	502	222	297	197

Table B.1 (cont.)

GNO[a]	$RAW_1 = 0.5, RAW_j$ = maximum, j = 2,3,4				$RAW_1 = 0.5$, i = 1,2,3,4			
	RAW_1	RAW_2	RAW_3	RAW_4	RAW_1	RAW_2	RAW_3	RAW_4
1	-135	252	-287	456	-174	241	105	165
2	94	177	-80	207	-174	230	206	77
3	359	53	233	147	46	88	286	75
4	-79	374	337	118	-22	277	247	58
5	1,030	87	366	-22	504	59	229	10
6	-392	199	533	224	-272	151	363	126
7	1,093	181	-435	146	624	244	157	108
8	-44	109	691	168	-217	117	428	66
9	161	260	-77	231	127	190	14	108
10	1,487	244	-285	42	735	237	-37	1
11	968	159	-41	240	366	136	24	136
12	617	281	-292	232	271	214	106	130
13	737	237	60	127	306	240	242	33
14	651	135	-135	191	175	193	50	128
15	176	69	-73	240	21	89	22	158
16	119	82	-134	158	-1	102	-59	122
17	-88	181	517	291	-175	164	412	90
18	327	302	-94	284	170	237	12	154
19	-111	254	362	168	-101	195	301	40
20	-413	172	116	300	-315	199	228	160
21	623	245	-576	230	339	290	213	138
22	-217	6	-179	238	-447	81	54	153

Source: Bailey (1988).

Note: Marginal products, in kg./ha., estimated for a change in RAW_i of 0.1. The estimated production functions are in Table 8.

[a]Numbers refer to genotypes. See Appendix A for descriptions.

Table C.1 Mean Yield, Regression Coefficients, and MSE from Stability Analysis, Actual Experiment Yields, Ranked by Mean Yield

GNO[a]	Mean Yield (kg./ha.)	a_g[b]	b_g[b]	MSE[c]
4	2,237	-44.4	1.26	64,095
18	2,101	62.1	1.12	61,696
14	2,052	-4.7	1.13	55,959
6	2,045	-56.3	1.16	76,530
17	2,040	171.8	1.03	108,988
11	1,980	-108.0	1.15	55,599
3	1,961	-42.7	1.10	65,453
13	1,954	-149.5	1.16	54,749
2	1,915	-117.8	1.12	84,354
12	1,852	-20.7	1.03	93,142
19	1,828	73.9	0.97	68,653
5	1,791	-25.8	1.00	81,866
9	1,741	248.1	0.82	44,369
10	1,711	106.7	0.88	83,146
20	1,666	-234.2	1.05	68,346
7	1,657	83.4	0.87	90,285
21	1,650	65.6	0.87	82,077
15	1,637	94.8	0.84	97,734
1	1,616	-61.7	0.92	48,819
8	1,595	-211.7	0.99	81,668
16	1,509	267.9	0.68	121,341
22	1,452	-96.8	0.85	75,742
Mean	1,818			

Source: Bailey (1988).

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]Regression coefficients from equation (1).

[c]Mean square error of the regression.

Table C.2 Mean Yield, Regression Coefficients, and MSE from Stability Analysis, Predicted Experiment Yields, Ranked by Mean Yield

GNO[a]	Mean Yield (kg./ha.)	a_g[b]	b_g[b]	MSE[c]
4	2,192	-96.8	1.29	11,165
18	2,046	115.6	1.09	18,225
14	2,009	-22.5	1.15	3,357
17	1,996	-89.8	1.18	30,977
6	1,992	-54.6	1.15	15,499
3	1,922	-22.8	1.10	7,220
11	1,916	-84.2	1.13	8,903
13	1,913	-42.6	1.10	13,850
2	1,873	-217.1	1.18	14,128
12	1,804	86.0	0.97	22,748
19	1,773	-30.3	1.02	10,765
5	1,740	173.3	0.88	24,496
9	1,705	336.6	0.77	11,554
10	1,689	170.1	0.86	36,784
21	1,633	226.8	0.79	21,370
7	1,623	266.9	0.76	35,310
20	1,622	-347.1	1.11	14,813
1	1,604	-178.8	1.00	14348
15	1,570	147.4	0.80	23,457
8	1,564	-377.1	1.09	23,591
16	1,436	257.7	0.66	10,676
22	1,429	-216.8	0.93	34,695
Mean	1,775			

Source: Bailey (1988).

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]Regression coefficients from equation (1).

[c]Mean square error of the regression.

Table C.3 Mean Yield, Regression Coefficients, and MSE from Stability Analysis, Simulated Yields, Hyderabad, Ranked by Mean Yield

GNO[a]	Mean Yield (kg./ha.)	a_g[b]	b_g[b]	MSE[c]
8	5,171	-1152.2	1.63	270,707
4	5,106	-145.2	1.36	21,701
17	5,014	-492.9	1.42	139,335
6	4,610	70.3	1.17	128,419
20	4,577	-589.7	1.34	57,156
11	4,545	-518.0	1.31	40,745
3	4,431	-174.9	1.19	8,101
18	4,354	117.5	1.09	24,403
14	4,347	-42.4	1.13	20,676
2	4,272	-187.6	1.15	14,382
19	4,169	-276.2	1.15	49,580
13	3,843	351.1	0.90	39,613
5	3,631	725.7	0.75	178,639
1	3,612	-153.5	0.97	25,739
15	3,451	-94.8	0.92	28,369
22	3,257	-334.8	0.93	76,966
12	3,255	271.5	0.77	55,963
9	3,004	449.2	0.66	28,047
10	2,767	560.6	0.57	97,575
16	2,737	214.6	0.65	22,904
21	2,568	686.3	0.49	110,329
7	2,444	715.6	0.45	75,557
Mean	3,871			

Source: Bailey (1988).

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]Regression coefficients from equation (1).

[c]Mean square error of the regression.

Table C.4 Mean Yield, Regression Coefficients, and MSE from Stability Analysis, Simulated Yields, Anantapur, Ranked by Mean Yield

GNO[a]	Mean Yield (kg./ha.)	a_g[b]	b_g[b]	MSE[c]
17	1,382	-158.3	1.53	481,715
6	1,330	-1.6	1.32	197,443
8	1,193	-271.4	1.46	436,446
3	1,170	-96.0	1.26	172,954
5	1,156	410.4	0.74	477,539
4	1,142	-57.4	1.19	93,756
14	1,064	-9.0	1.07	45,204
18	1,059	89.9	0.96	121,946
19	1,054	-89.3	1.14	113,575
2	1,035	-138.6	1.17	79,930
20	1,012	-237.8	1.24	187,786
9	975	223.3	0.75	126,244
15	971	135.5	0.83	88,759
13	967	-163.6	1.12	44,698
1	964	-89.3	1.05	150,267
11	935	6.4	0.92	90,113
7	843	93.6	0.75	253,668
21	827	117.4	0.70	370,361
16	813	216.8	0.59	84,378
22	764	-64.4	0.82	172,720
10	742	13.4	0.72	313,196
12	727	70.0	0.65	158,382
Mean	1,006			

Source: Bailey (1988).

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]Regression coefficients from equation (1).

[c]Mean square error of the regression.

Table C.5. Mean Yield, Regression Coefficients, and MSE from Stability Analysis, Simulated Yields, Gujarat, Ranked by Mean Yield

GNO[a]	Mean Yield (kg./ha.)	a_g [b]	b_g [b]	MSE[c]
8	4,171	-337.0	1.48	83,931
17	4,035	-94.4	1.36	40,349
4	3,984	-67.3	1.33	87,648
20	3,728	-124.8	1.27	207,393
6	3,679	194.7	1.14	78,424
11	3,540	-210.3	1.23	136,605
3	3,535	-79.6	1.19	51,801
14	3,458	87.8	1.11	3,689
2	3,405	-148.9	1.17	154,149
18	3,373	87.3	1.08	98,948
19	3,297	-119.4	1.12	8,769
5	3,082	894.9	0.72	419,569
13	2,917	-153.7	1.01	95,942
1	2,760	-362.6	1.03	185,181
15	2,700	-52.2	0.90	5,618
22	2,539	-343.0	0.95	40,752
12	2,452	-30.5	0.82	136,373
9	2,236	8.2	0.73	61,903
16	2,200	286.7	0.63	26,095
10	2,142	193.3	0.64	31,781
21	1,909	116.8	0.59	30,413
7	1,889	253.7	0.54	43,929
Mean	3,047			

Source: Bailey (1988).

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]Regression coefficients from equation (1).

[c]Mean square error of the regression.

Table D.1 Mean Yields and Standard Deviations, Ranked by Yield, Predicted Experimental Yields, and Simulated Yields for Hyderabad, Anantapur and Gujarat

Experiment			Hyderabad			Anantapur			Gujarat		
GNO[a]	Mean	Std.	GNO[a]	Mean	Std.	GNO[a]	Mean	Std.	GNO[a]	Mean	Std.
4	2192	1127	8	5171	2046	17	1382	1715	8	4171	2895
18	2046	956	4	5106	1650	6	1330	1428	17	4035	2646
14	2009	998	17	5013	1763	8	1192	1631	4	3983	2605
17	1996	1038	6	4610	1465	3	1170	1355	20	3728	2503
6	1992	1011	20	4577	1634	5	1156	1024	6	3679	2243
3	1922	958	11	4545	1597	4	1142	1260	11	3540	2424
11	1916	986	3	4431	1444	14	1064	1115	3	3535	2320
13	1913	966	18	4354	1335	18	1059	1048	14	3458	2154
2	1873	1032	14	4347	1381	19	1055	1214	2	3405	2304
12	1804	856	2	4272	1401	2	1034	1229	18	3373	2122
19	1773	891	19	4169	1409	20	1012	1345	19	3297	2185
5	1739	784	13	3843	1111	9	975	843	5	3082	1536
9	1705	679	5	3631	1002	15	971	901	13	2917	1986
10	1689	769	1	3612	1189	13	967	1172	1	2760	2040
21	1633	705	15	3451	1122	1	964	1141	15	2700	1760
7	1623	691	22	3257	1158	11	935	993	22	2539	1852
20	1622	973	12	3255	963	7	843	913	12	2452	1628
1	1604	882	9	3004	817	21	827	942	9	2236	1445
15	1570	714	10	2767	757	16	813	673	16	2200	1233
8	1564	964	16	2737	804	22	764	941	10	2142	1258
16	1436	587	21	2568	675	10	742	928	21	1909	1158
22	1429	828	7	2444	606	12	727	778	7	1889	1065

Source: Bailey (1988).

[a]Numbers refer to genotypes. See Appendix A for descriptions.

Table D.2 Gini Coefficients, Predicted Yields from Trial, Ranked by Mean Yield

GNO[a]	μ	Γ	$\mu - \Gamma$
4	2,191.63	649.90	1,541.73
18	2,045.91	550.08	1,495.82
14	2,009.03	575.96	1,433.07
17	1,995.91	599.54	1,396.37
6	1,991.91	584.96	1,406.95
3	1,922.32	552.49	1,369.83
11	1,916.38	570.89	1,345.49
13	1,913.36	556.00	1,357.36
2	1,872.50	596.08	1,276.42
12	1,803.50	492.13	1,311.38
19	1,773.37	514.52	1,258.85
5	1,739.47	450.93	1,288.54
9	1,704.53	390.38	1,314.15
10	1,688.82	438.98	1,249.83
21	1,632.98	400.12	1,232.86
7	1,622.74	387.85	1,234.89
20	1,621.58	563.11	1,058.47
1	1,603.64	510.41	1,093.23
15	1,570.05	410.08	1,159.97
8	1,564.12	557.63	1,006.50
16	1,435.84	336.12	1,099.72
22	1,428.60	479.64	948.96

MΓ[b] efficient set: 4, 18, 12, 5, 9, 7, 16

MG[c] efficient set: 4

Source: Bailey 1988

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]The MΓ criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\Gamma_F \leq \Gamma_G$, with at least one strict inequality.

[c]The MG criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\mu_F - \Gamma_F \geq \mu_G - \Gamma_G$, with at least one strict inequality (Yitzhaki, 1982).

Table D.3 Gini Coefficients, Simulated Yields for Hyderabad, Ranked by Mean Yield

GNO[a]	μ	Γ	μ - Γ
8	5,171.34	1061.23	4,110.10
4	5,106.18	826.88	4,279.30
17	5,013.48	893.10	4,120.39
6	4,610.31	734.74	3,875.57
20	4,576.83	838.90	3,737.93
11	4,545.24	811.23	3,734.00
3	4,431.22	725.93	3,705.29
18	4,353.85	656.97	3,696.88
14	4,346.98	686.64	3,660.34
2	4,272.37	702.36	3,570.01
19	4,168.99	708.69	3,460.30
13	3,843.00	520.88	3,322.12
5	3,631.23	493.29	3,137.95
1	3,611.73	584.62	3,027.10
15	3,450.60	558.15	2,892.45
22	3,257.39	588.14	2,669.25
12	3,254.74	459.97	2,794.77
9	3,004.38	375.88	2,628.50
10	2,767.35	342.68	2,424.67
16	2,737.47	394.19	2,343.28
21	2,567.51	281.08	2,286.43
7	2,444.13	254.60	2,189.53

MΓ[b] efficient set: 8, 4, 6, 3, 18, 13, 5, 12, 9, 10, 21, 7

MG[c] efficient set: 8, 4

Source: Bailey 1988

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]The MΓ criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\Gamma_F \leq \Gamma_G$, with at least one strict inequality.

[c]The MG criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\mu_F - \Gamma_F \geq \mu_G - \Gamma_G$, with at least one strict inequality (Yitzhaki, 1982).

Table D.4 Gini Coefficients, Simulated Yields for Anantapur Ranked by Mean Yield

GNO[a]	μ	Γ	μ - Γ
17	1,381.79	869.43	512.36
6	1,330.10	753.71	576.39
8	1,192.47	790.93	401.55
3	1,170.00	704.88	465.13
5	1,155.58	553.40	602.18
4	1,142.10	655.21	486.89
14	1,064.23	594.81	469.43
18	1,059.31	554.77	504.55
19	1,054.85	630.38	424.47
2	1,034.45	638.69	395.77
20	1,011.56	666.05	345.51
9	974.61	465.39	509.22
15	970.58	495.53	475.05
13	967.20	604.77	362.43
1	963.76	596.72	367.04
11	935.10	498.93	436.17
7	842.72	481.76	360.95
21	826.53	487.84	338.68
16	812.73	374.64	438.10
22	764.39	482.03	282.36
10	742.05	460.18	281.87
12	727.44	401.33	326.11

MΓ[b] efficient set: 17, 6, 3, 5, 9, 16

MG[c] efficient set: 17, 6, 5

Source: Bailey 1988

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]The MΓ criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\Gamma_F \leq \Gamma_G$, with at least one strict inequality.

[c]The MG criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\mu_F - \Gamma_F \geq \mu_G - \Gamma_G$, with at least one strict inequality (Yitzhaki, 1982).

Table D.5 Gini Coefficients, Simulated Yields for Gujarat Ranked by Mean Yield

GNO[a]	μ	Γ	μ - Γ
8	4,171.41	577.85	3,593.56
17	4,035.04	518.67	3,516.37
4	3,983.46	517.46	3,466.00
20	3,728.42	498.86	3,229.57
6	3,678.74	449.37	3,229.37
11	3,539.97	481.26	3,058.71
3	3,534.79	455.40	3,079.40
14	3,457.63	425.87	3,031.75
2	3,405.13	456.19	2,948.94
18	3,372.78	422.03	2,950.75
19	3,296.72	431.04	2,865.67
5	3,082.37	303.28	2,779.09
13	2,916.67	391.14	2,525.53
1	2,760.01	404.28	2,355.73
15	2,700.21	347.69	2,352.53
22	2,538.68	362.80	2,175.87
12	2,452.33	323.72	2,128.61
9	2,236.29	286.73	1,949.56
16	2,200.38	246.12	1,954.25
10	2,142.27	244.79	1,897.47
21	1,909.37	227.71	1,681.66
7	1,889.09	208.72	1,680.37

MΓ[b] efficient set: 8, 17, 4, 20, 6, 14, 18, 5, 9, 16, 10, 21, 7

MG[c] efficient set: 8

Source: Bailey 1988

[a]Numbers refer to genotypes. See Appendix A for descriptions.

[b]The MΓ criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\Gamma_F \leq \Gamma_G$, with at least one strict inequality.

[c]The MG criterion is: F dominates G if $\mu_F \geq \mu_G$ and $\mu_F - \Gamma_F \geq \mu_G - \Gamma_G$, with at least one strict inequality (Yitzhaki, 1982).

Table D.6 Results from EUMGF Analysis, Five Highest Ranked Genotypes

Location	Values of r_A											
	0.0[a]		0.00005		0.00015		0.00037		0.00198		0.005	
	GNO[b]	CE[c]	GNO[b]	CE[c]	GNO[b]	CE[c]	GNO[b]	CE[c]	GNO[b]	CE[c]	GNO[b]	CE[c]
Experiment	4	2,192	4	2,160	4	2,097	4	1,957	18	1,271	9	876
	18	2,046	18	2,023	18	1,977	18	1,876	9	1,260	5	773
	14	2,009	14	1,984	14	1,935	14	1,824	4	1,214	18	770
	17	1,996	17	1,969	17	1,916	6	1,804	14	1,181	11	739
	6	1,992	6	1,967	6	1,916	17	1,801	17	1,177	16	734
Hyderabad	8	5,171	8	5,065	4	4,881	4	4,464	6	1,830	5	747
	4	5,106	4	5,036	8	4,833	17	4,309	5	1,826	6	747
	17	5,013	17	4,934	17	4,759	8	4,269	4	1,785	8	746
	6	4,610	6	4,556	6	4,436	6	4,124	18	1,775	20	746
	20	4,577	20	4,509	20	4,361	20	3,989	17	1,773	17	745
Anantapur	17	1,382	17	1,312	6	1,191	6	1,026	5	534	9	269
	6	1,330	6	1,281	17	1,187	5	984	9	512	5	265
	8	1,192	5	1,130	5	1,081	17	975	6	504	6	263
	3	1,170	8	1,130	3	1,045	4	905	18	490	18	252
	5	1,156	3	1,126	4	1,034	3	896	15	471	16	251
Gujarat	8	4,171	8	3,966	8	3,553	17	2,812	5	1,326	4	789
	17	4,035	17	3,863	17	3,515	4	2,799	6	1,249	5	776
	4	3,983	4	3,817	4	3,480	6	2,772	18	1,199	6	766
	20	3,728	20	3,575	6	3,303	8	2,749	4	1,198	3	706
	6	3,679	6	3,555	20	3,266	20	2,642	14	1,178	18	690

Source: Bailey 1988.

[a] r_A = 0.0 approximated by 0.00000001.

[b] GNO refers to genotypes described in Appendix A.

[c] CE = certainty equivalent (kg./ha.) calculated from equation (19) and rounded to nearest integer.

Table D.7 Extended Gini Coefficients, Predicted Yields from Trial, Ranked by Mean Yield[a]

GNO[b]	μ	v=3		v=5		v=8	
		$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$
4	2,192	985	1,207	1,336	855	1,572	620
18	2,046	837	1,209	1,143	903	1,356	690
14	2,009	873	1,136	1,190	819	1,413	596
17	1,996	891	1,105	1,186	810	1,389	607
6	1,992	882	1,110	1,196	796	1,410	582
3	1,922	839	1,083	1,146	776	1,361	561
11	1,916	853	1,064	1,140	776	1,331	586
13	1,913	851	1,063	1,171	743	1,396	517
2	1,873	894	978	1,201	672	1,404	468
12	1,804	752	1,051	1,034	769	1,236	567
19	1,773	771	1,002	1,037	736	1,220	553
5	1,739	691	1,049	955	784	1,143	597
9	1,705	598	1,107	827	878	996	709
10	1,689	674	1,015	943	746	1,147	542
21	1,633	622	1,011	876	757	1,068	565
7	1,623	606	1,017	867	755	1,076	546
20	1,622	839	782	1,116	505	1,292	330
1	1,604	761	843	1,015	589	1,186	418
15	1,570	628	942	864	706	1,031	539
8	1,564	818	746	1,070	494	1,228	336
16	1,436	517	919	717	719	864	572
22	1,429	714	715	950	478	1,105	324

GNO[b]	μ	v=10		v=20		v=50	
		$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$
4	2,192	1,659	533	1,861	331	2,018	173
18	2,046	1,438	608	1,642	404	1,814	232
14	2,009	1,500	509	1,710	299	1,874	135
17	1,996	1,470	526	1,683	312	1,865	131
6	1,992	1,490	501	1,672	320	1,796	196
3	1,922	1,443	479	1,638	284	1,786	136
11	1,916	1,401	515	1,567	350	1,698	218
13	1,913	1,482	431	1,681	232	1,814	99
2	1,873	1,481	392	1,660	212	1,786	86
12	1,804	1,315	489	1,505	298	1,646	157
19	1,773	1,292	481	1,476	297	1,637	137
5	1,739	1,213	526	1,368	371	1,459	281
9	1,705	1,063	641	1,227	478	1,341	364
10	1,689	1,230	459	1,432	257	1,580	109
21	1,633	1,146	487	1,345	288	1,511	122
7	1,623	1,163	460	1,376	247	1,518	105
20	1,622	1,354	268	1,486	136	1,559	62
1	1,604	1,251	353	1,408	196	1,524	80
15	1,570	1,096	474	1,255	315	1,395	175
8	1,564	1,285	280	1,410	154	1,489	75
16	1,436	923	513	1,080	356	1,229	206
22	1,429	1,161	268	1,288	140	1,370	59

[a]Calculation based on equation (12).

[b]GNO refers to genotypes described in Appendix A.

Table D.8 Extended Gini Coefficients, Simulated Yields for Hyderabad, Ranked by Mean Yield[a]

GNO[b]	μ	v=3 $\Gamma(v)$	v=3 $\mu-\Gamma(v)$	v=5 $\Gamma(v)$	v=5 $\mu-\Gamma(v)$	v=8 $\Gamma(v)$	v=8 $\mu-\Gamma(v)$
8	5,171	1,777	3,395	2,648	2,524	3,316	1,856
4	5,106	1,407	3,699	2,160	2,946	2,812	2,294
17	5,013	1,514	3,499	2,305	2,708	2,959	2,055
6	4,610	1,249	3,361	1,914	2,697	2,476	2,134
20	4,577	1,415	3,162	2,137	2,439	2,725	1,852
11	4,545	1,371	3,175	2,078	2,467	2,665	1,880
3	4,431	1,234	3,197	1,890	2,541	2,452	1,979
18	4,354	1,124	3,230	1,741	2,613	2,284	2,070
14	4,347	1,171	3,176	1,803	2,544	2,351	1,996
2	4,272	1,193	3,079	1,828	2,445	2,375	1,897
19	4,169	1,204	2,965	1,844	2,325	2,388	1,781
13	3,843	9,04	2,939	1,435	2,408	1,932	1,911
5	3,631	8,39	2,792	1,294	2,337	1,692	1,939
1	3,612	1,003	2,609	1,557	2,054	2,048	1,564
15	3,451	952	2,498	1,465	1,986	1,907	1,544
22	3,257	992	2,266	1,500	1,757	1,925	1,332
12	3,255	794	2,461	1,248	2,007	1,670	1,584
9	3,004	654	2,350	1,043	1,961	1,417	1,588
10	2,767	596	2,171	966	1,801	1,334	1,433
16	2,737	675	2,063	1,046	1,691	1,378	1,360
21	2,568	501	2,066	837	1,730	1,184	1,383
7	2,444	451	1,993	750	1,694	1,062	1,382

GNO[b]	μ	v=10 $\Gamma(v)$	v=10 $\mu-\Gamma(v)$	v=20 $\Gamma(v)$	v=20 $\mu-\Gamma(v)$	v=50 $\Gamma(v)$	v=50 $\mu-\Gamma(v)$
8	5,171	3,577	1,594	4,215	956	4,759	413
4	5,106	3,099	2,007	3,906	1,200	4,665	442
17	5,013	3,233	1,780	3,961	1,053	4,607	407
6	4,610	2,718	1,892	3,392	1,218	4,085	525
20	4,577	2,969	1,608	3,609	968	4,179	397
11	4,545	2,916	1,630	3,606	939	4,222	323
3	4,431	2,699	1,733	3,392	1,040	4,042	389
18	4,354	2,528	1,826	3,237	1,117	3,942	412
14	4,347	2,594	1,753	3,287	1,060	3,954	393
2	4,272	2,619	1,654	3,312	960	3,944	329
19	4,169	2,623	1,546	3,267	902	3,843	326
13	3,843	2,164	1,679	2,854	989	3,524	319
5	3,631	1,870	1,761	2,404	1,227	3,069	562
1	3,612	2,266	1,345	2,877	735	3,397	215
15	3,451	2,101	1,350	2,653	798	3,172	278
22	3,257	2,110	1,147	2,631	626	3,076	182
12	3,255	1,868	1,387	2,455	800	3,007	248
9	3,004	1,596	1,409	2,150	854	2,722	283
10	2,767	1,511	1,256	2,037	730	2,523	244
16	2,737	1,529	1,209	1,978	760	2,446	291
21	2,568	1,355	1,213	1,871	697	2,354	213
7	2,444	1,217	1,227	1,698	746	2,187	257

[a]Calculation based on equation (12).

[b]GNO refers to genotypes described in Appendix A.

Table D.9 Extended Gini Coefficients, Simulated Yields for Anantapur, Ranked by Mean Yield[a]

GNO[b]		v=3		v=5		v=8		v=10	
	μ	Γ(v)	μ-Γ(v)	Γ(v)	μ-Γ(v)	Γ(v)	μ-Γ(v)	Γ(v)	μ-Γ(v)
17	1,382	1,136	245	1,301	81	1,360	21	1,372	10
6	1,330	1,016	314	1,204	126	1,287	43	1,306	24
8	1,192	1,010	182	1,136	57	1,177	16	1,185	8
3	1,170	937	233	1,091	79	1,148	22	1,159	11
5	1,156	793	362	1,004	151	1,106	49	1,130	26
4	1,142	882	260	1,045	97	1,112	30	1,126	16
14	1,064	815	249	979	85	1,043	21	1,055	10
18	1,059	767	292	941	118	1,021	38	1,040	19
19	1,055	841	214	981	74	1,034	21	1,045	10
2	1,034	851	184	983	52	1,024	11	1,030	5
20	1,012	860	152	971	41	1,003	8	1,008	4
9	975	662	312	836	139	925	50	948	27
15	971	698	273	864	107	939	32	955	15
13	967	800	167	920	47	957	10	963	5
1	964	796	168	917	47	954	10	959	4
11	935	682	253	832	104	902	33	918	17
7	843	654	189	780	62	828	15	836	6
21	827	653	174	769	58	813	14	821	6
16	813	544	269	698	115	775	38	794	19
22	764	635	129	728	36	757	8	761	3
10	742	602	140	696	46	731	11	737	5
12	727	542	185	653	75	704	24	716	12

[a]Calculations based on equation (12).

[b]GNO refers to genotypes described in Appendix A.

Table D.10 Extended Gini Coefficients, Simulated Yields for Gujarat, Ranked by Mean Yield[a]

GNO[b]		v=3		v=5		v=8		v=10		v=20	
	μ	$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$	$\Gamma(v)$	$\mu-\Gamma(v)$
8	4,171	2,403	1,769	3,254	917	3,675	496	3,776	395	3,889	282
17	4,035	2,170	1,865	2,959	1,076	3,354	681	3,453	582	3,610	425
4	3,983	2,154	1,829	2,921	1,063	3,299	684	3,390	594	3,504	480
20	3,728	2,065	1,664	2,806	923	3,190	539	3,292	436	3,452	277
6	3,679	1,887	1,792	2,600	1,079	2,989	690	3,096	583	3,252	427
11	3,540	1,988	1,552	2,669	871	2,992	548	3,069	471	3,175	365
3	3,535	1,911	1,624	2,619	916	2,972	563	3,054	480	3,142	393
14	3,458	1,785	1,673	2,445	1,012	2,794	663	2,890	567	3,064	393
2	3,405	1,891	1,514	2,570	835	2,909	496	2,993	413	3,112	294
18	3,373	1,751	1,622	2,370	1,003	2,687	686	2,773	599	2,937	436
19	3,297	1,802	1,494	2,456	841	2,778	518	2,855	441	2,955	342
5	3,082	1,316	1,766	1,909	1,173	2,305	777	2,436	646	2,657	425
13	2,917	1,633	1,283	2,236	681	2,539	378	2,609	307	2,691	226
1	2,760	1,665	1,095	2,243	517	2,514	246	2,572	188	2,622	138
15	2,700	1,450	1,251	1,968	732	2,224	477	2,286	414	2,382	319
22	2,539	1,515	1,024	2,057	482	2,313	226	2,367	171	2,413	125
12	2,452	1,350	1,103	1,834	618	2,079	374	2,141	312	2,233	219
9	2,236	1,196	1,041	1,625	611	1,840	397	1,893	343	1,972	264
16	2,200	1,029	1,171	1,413	788	1,628	572	1,696	505	1,845	355
10	2,142	1,040	1,102	1,453	689	1,683	459	1,748	395	1,861	281
21	1,909	965	944	1,344	565	1,556	354	1,616	294	1,719	191
7	1,889	883	1,006	1,237	652	1,441	448	1,501	389	1,614	275

[a]Calculations based on equation (12).

[b]GNO refers to genotypes described in Appendix A.

APPENDIX E

Determination of Intervals of Absolute Risk Aversion

If upper and lower bounds on the absolute risk aversion coefficient, $r_A(x) = -U''(x)/U'(x)$, can be specified then SDWRF can be used to order genotypes for the different sub-groups of decision makers represented by those intervals. Such intervals on r_A can be derived from the results reported by Binswanger (1978, 1980) of games where Indian farmers were asked to choose between a number of alternative gambles each with two payoffs of equal probability decided by the toss of a coin as follows:

Game	Payoff		Risk Aversion Class
	Heads	Tails	
F	0	200	Neutral to negative
E	10	190	Slight to neutral
C	30	150	Moderate
B	40	120	Intermediate
A	45	95	Severe
O	50	50	Extreme

Risk aversion is measured by the partial risk aversion function:

(E.1) $P(W; x) = -x\,[U''(W+x)/U'(W+x)]$

where W = initial wealth, x = a gain or loss from a gamble, and $w = W+x$ = total assets = W + gain, or W - loss (Menezes and Hanson 1970; Zeckhauser and Keeler 1970). $P(W; x)$ is simply a multiple of the absolute risk aversion function of total assets (Binswanger 1978):[1]

(E.2) $P(W; x) = xr_A(W+x)$.

Given the games' payoffs, Binswanger estimates P by approximating it on a utility function with constant partial risk aversion:

(E.3) $U(M) = (1-P)M^{1-P}$

where M is the certainty equivalent of a gamble such that $U(M)=EU$. If X_{ij} is the jth outcome of gamble i ($i = 1,2$; $j = 1,2$) and all X_{ij} have probability of 1/2, then M and P can be estimated sequentially from the equations (Binswanger 1978):

(E.4) $U(M) = 1/2\,(1-P)[X_{11}^{1-P}+X_{12}^{1-P}] = 1/2\,(1-P)[X_{21}^{1-P}+X_{22}^{1-P}]$

---

[1] Note that the partial risk aversion function should not be confused with Pratt's relative risk aversion function: $r_R(x) = -x[U''(x)/U'(x)] = xr_A(x)$.

Binswanger (1978; 1980) presents the bounds on the values of P implied by the choices in the game. From these bounds one can derive M, the certainty equivalent at the point of indifference between two gambles,

(E.5) $M = U^{-1}EU = [1/2\ (X_{i1}^{1-P} + X_{i2}^{1-P})]^{1/1-P}$

and the corresponding bounds on the absolute risk aversion coefficient as:

(E.6) $r_A = P/M$

Binswanger's reported bounds on P, and our estimated M and r_A, are presented in Table E.1.

Table E.1. Partial and Absolute Risk Aversion Coefficients for Binswanger's Risk Aversion Classes

Game	Payoff		Risk Aversion	P[a]		M[b]	r_A[c]	
	Heads	Tails	Class	lower bound	upper bound		lower bound	upper bound
	0	200	Neutral to Negative	$-\infty$	0		$-\infty$	0
						100		
E	10	190	Slight to Neutral	0	0.315		0	0.0038
						82.89		
C	30	150	Moderate	0.315	0.812		0.0038	0.0114
						71.28		
B	40	120	Intermediate	0.812	1.74		0.0114	0.0280
						62.15		
A	45	95	Severe	1.74	7.51		0.0280	0.1502
						50		
O	50	50	Extreme	7.51	∞		0.1502	∞

[a] Binswanger (1980)

[b] Certainty equivalent at point of indifference between two games.

[c] Estimated $r_A = P/M$.

Partial risk aversion, $P(W; x)$, is fixed regardless of the level of the payoff (Zeckhauser and Keeler 1970; Menezes and Hanson 1970). However, r_A is affected by a multiplicative transformation of the payoffs of the game and in order to use the bounds of r_A estimated from the payoffs in the game, they must be transformed so that they correspond to payoffs in terms of kg./ha. of groundnuts. Raskin and Cochran address this problem: if $r_A(x)$ is elicited as $r/\$$, as a measure of aversion to annual income risk, for instance, and if the crop under consideration represents the entire income of the farmer and the farm area is known, then the conversion factor between the two scales would be the area, A, and $r_A(w) = Ar/\$$, where w is in per unit area income dollars (Raskin and Cochran 1986, p.207).

A number of generalizations are made in order to transform Binswanger's results into meaningful bounds on r_A for the groundnut genotype yield distributions. Of the six villages included in Binswanger's study, groundnuts are an important crop, in terms of percent gross area, in one (Dokur) and a secondary crop in two others (Binswanger 1978; Jodha 1977). Rescaling of r_A is based on the average area sown to groundnuts, in both sole and mixed crops, in Dokur, over a sample of small, medium and large farmers, estimated from data presented by Jodha (1977) as 0.56 ha. The average product price for groundnuts is 400Rs./quintal = 4Rs./kg. (Tom Walker, personal communication). Yields, in kg./ha., can thus be converted into a gross income from an average area of A = 0.56 ha.; if yield = y kg./ha., then gross income = 4Ay Rs. = 2.24y Rs. Yields of genotypes range from a minimum of zero in severely droughted years (treatments) to a maximum of 6800 kg/ha under non-stressed conditions, equal to approximately 15,200 Rs. gross income from 0.56ha, which is approximately 76 times the maximum payoff in the game. Rescaling the estimated r_A from the game therefore involves a conversion factor of 4(0.56)/76 = 0.02945. The bounds of r_A thus obtained and used in the SDWRF analysis are given in Table E.2. Arbitrary bounds, of -0.001 and 0.005, were set at the upper and lower limits. The full range of r_A for risk averse decision makers (0, 0.005), is included as a subset of the conditions for SSD.

Table E.2. Estimated Bounds on Risk Aversion Intervals used in Generalized Stochastic Dominance (SDWRF) Analysis

Risk Aversion Class	Bound on r_A		Code
	lower	upper	
Neutral to negative	-0.0001	0	RNEG
Neutral	0	0	R0
Slight to neutral	0	0.00005	R1
Moderate	0.00005	0.00015	R2
Intermediate	0.00015	0.00037	R3
Severe	0.00037	0.00198	R4
Extreme	0.00198	0.005	R5
Risk Averse	0	0.005	RALL

Other Agricultural Economics Research Papers

No. 88-16	A Progress Report on the New York FarmNet Program, April 1, 1987-March 31, 1988	C. Delaney
No. 88-17	Consumer Segmentation Analysis of Grocery Coupon Users	M. Meloy E. McLaughlin C. Kramer
No. 89-1	The Competitiveness of New York State Onions During the 1987-88 Marketing Year	E. Figueroa
No. 89-2	An Analysis of the Acceptance of IPM Techniques in Processed Sweet Corn	J. Waldorph G. B. White
No. 89-3	Statistical Summary of the 1987 Farm Management and Energy Survey	Michael J. Kelleher Nelson L. Bills
No. 89-4	Microcomputers and Small Local Governments in New York: Five Case Studies	Duane E. Wilcox
No. 89-5	Farmer Investment Behavior: A Review of Literature	Brian T. Brase Eddy L. LaDue
No. 89-6	Eighty Years of Change in Dairy Farming Dryden Township, Tompkins County, New York	B. F. Stanton
No. 89-7	1988 Budget Guide, Estimated Prices for Crop Operating Inputs and Capital Investment Items	Darwin Snyder
No. 89-8	A Microcomputer Program for Projecting Annual Cash Flows, Debt Repayment Ability and Proforma Financial Statements	Eddy L. LaDue David B. Cook
No. 89-9	Diversification of the Cheddar Cheese Industry Through Specialty Cheese Production: An Economic Assessment	John C. Martin David M. Barbano Richard D. Aplin
No. 89-10	Management Strategies to Improve Profitability on Limited Resource Dairy Farms: A Linear Programming Analysis	B. F. Stanton