Theor Appl Genet (1991) 81: 825–832

THEORETICAL AND APPLIED GENETICS

Expected utility maximization and selection of stable plant cultivars

K. M. Eskridge [1,*] **and B. E. Johnson** [2]

[1] Department of Biometry, and [2] Department of Agronomy, University of Nebraska – Lincoln, Lincoln, NE 68583-0712, USA

Received November 15, 1989; Accepted July 13, 1990
Communicated by A. R. Hallauer

Summary. In most plant breeding programs, selection of the best commercially suitable cultivars for a target group of environments is based on information obtained from evaluation trials cultivated in a sample of environments. Information on the performance of cultivars collected in a sample of environments can only be approximate and, consequently, selection of the best cultivar involves choosing among cultivars that respond uncertainly in many environments. The agronomic and/or economic value of a cultivar across environments may be considered the general or overall utility of the cultivar. Data from a sample of environments therefore provides only an estimate of any cultivar's overall utility, with the overall goal of selection among all cultivars being the maximization of the expected utility. Within this framework, expected utility maximization, an approach to decision making that has been well developed in the disciplines of economics and statistics, can assist the plant breeder in making such decisions. This research was initiated (1) to determine how expected utility maximization might be used to develop indices that are useful for selecting broadly adapted plant cultivars, and (2) to determine how the breeder's preferences might affect choice of the best cultivar. The data used in this research were from USDA Regional Soybean Tests. The results indicated that expected utility maximization, which explicitly incorporates into the selection rule the plant breeder's preferences regarding stability, can be a useful aid in the selection of stable plant cultivars.

Key words: Selection – Utility functions – Plant breeding – Genotype × environment interaction – Yield stability

* To whom correspondence should be addressed

Introduction

An important consideration in many plant breeding programs is the selection of cultivars that perform well over a wide range of environments. Selection of such broadly adapted cultivars is difficult, however, since the phenotypic response to changes in environment usually differs among cultivars. A multitude of univariate stability measures have been proposed to aid the plant breeder in identifying broadly adapted plant cultivars (Lin et al. 1986). Most of these approaches to stability have been developed under the implicit assumption that the breeder would use both the mean yield and stability when making selections. Yet little has been done to explicitly indicate how an index based on both mean yield and stability might be developed to aid the plant breeder in making selections. Consequently, plant breeders have been left to their own initiative in weighing the importance of stability relative to yield and in making final choices. If stability may be thought of as simply a measure of variability or uncertainty, then techniques for making decisions under risk can be used to develop 'stability indices,' which weigh the importance of mean yield relative to stability.

One approach to making decisions under risk that might be applied to plant selection is based on the assumption that the breeder chooses cultivars which have the largest expected utility, i.e., he/she practices expected utility maximization (EUM). EUM was developed to model how decision makers choose between alternatives that have uncertain outcomes (von Neumann and Morgenstern 1972; Anderson et al. 1977). In terms of plant selection, this theory asserts that the plant breeder who wants to maximize expected utility will judge the total value of a cultivar using a utility function, $U(Y)$, which associates a single real number to any value of yield Y.

Then, given any two cultivars, A and B, cultivar A will be preferred to cultivar B only if $E[U(Y_A)] > E[U(Y_B)]$, where E denotes the mathematical expectation operator. That is, cultivar A is preferred to cultivar B if the average (or expected) value of utility over all possible yield values is larger for cultivar A than for cultivar B.

The major importance of EUM behavior to plant breeding is that models resulting from such behavior may be used to develop indices that explicitly quantify how the EUM plant breeder weighs the importance of yield relative to stability, when developing cultivars for a broad range of environments. Here, a specific form of utility function was used to develop a selection index that illustrates how the EUM plant breeder weighs the importance of mean yield to stability. Using this specific form of the utility function and assuming stability may be thought of as a measure of variance; then the EUM plant breeder will prefer cultivars with large values of:

$$\bar{Y}_{i\cdot} - (a/2)\,(V_i)$$

where $\bar{Y}_{i\cdot}$ is the sample mean yield over environments for cultivar i, V_i is some variance measure of stability for cultivar i and "a" is a weight indicating the importance the breeder places on stability.

A limited amount of research has applied EUM to selection. Barah et al. (1981) used E-V analysis (a certain kind of EUM model) to order sorghum varieties. Schneeberger et al. (1982), Smith and Allaire (1985), and Smith and Hammond (1986) utilized EUM concepts in mate selection in animal breeding. Allaire and Thraen (1985) discussed the use of EUM concepts as a means of improving the economic efficiency of dairy cattle. None of these papers, however, showed how commonly used measures of stability might be included in the model or how different levels of concern about stability might affect the ordering of the cultivars. The purpose of this research was: 1) to determine how EUM concepts might be used to develop indices that are useful for selecting broadly adapted plant cultivars; and 2) to determine how the breeder's preferences, with regard to stability, might affect the ordering of the cultivars when using such indices.

Major components of expected utility maximization

Expected utility maximization applied to plant selection can be separated into four major steps: (1) enumeration of all possible choices; (2) valuation of the cultivars; (3) specification of probability distributions as 'predictions' of cultivar response; and (4) identification and use of a choice criterion to select the most valuable cultivar.

Enumeration of choices

The first step is to list all the possible choices available to the plant breeder. If the breeder is concerned with selecting the single 'best' cultivar from a set of cultivars, as will be assumed here, then the list of all possible choices is simply the set of cultivars being evaluated.

Valuation of cultivars

The breeder must be able to specify the 'value' associated with each cultivar. This 'value' might be stated in terms of yield, by an index of economic worth, profit per hectare, etc. However, these or similar measures of value may not be good reflections of the cultivar's value, primarily because they fail to consider the breeder's attitude toward stability. For example, two cultivars might have identical mean yields (economic worth, etc.), but if cultivar A is half as stable over environments as cultivar B, then cultivar B would generally be preferred.

A more general way of measuring cultivar value, which includes the breeder's attitude toward stability, is to use utility functions. A breeder's utility function can be defined as a subjective valuing mechanism reflecting the breeder's assessment of the total value of a cultivar as a function of a single trait, such as yield, or of a phenotypic index of agronomic and/or economic value. Research has shown that if the plant breeder (or any person for that matter) adheres to a set of reasonable behavioral axioms, then the mechanism that he/she uses to value cultivars can be characterized by a utility function (von Neumann and Morgenstern 1972; Anderson et al. 1977). (Note: in the following, all discussion will be in reference to yield; however, the method is entirely general and can be applied to any phenotypic measure of value, such as an index of exonomic or agronomic worth, merit, profit, etc.

There are several important considerations when using utility functions to evaluate cultivars. First, if the breeder prefers cultivar A to cultivar B, then the utility which the breeder assigns to cultivar A is larger than that of cultivar B and vice versa. Second, the scale on which utility is defined is arbitrary, analogous to a temperature scale. This means that the ordering of the cultivars must not change under a postive linear transformation. For example, $U^*(Y)$ will produce the same cultivar orderings as $U(Y)$ if $U^*(Y) = b_0 + b_1 U(Y)$, where $b_1 > 0$. Finally, a breeder's utility function is likely to be a concave function of yield, where the curvature of the utility function defines the breeder's attitude toward yield stability (Fig. 1). The more curved a breeder's utility function, the greater the importance he/she places on yield stability. This last result, called decreasing marginal utility of yield, is directly analogous to the concave utility function of a decision maker who is adverse to risk (Anderson et al. 1977), and it is discussed below.

Only a few studies have been conducted to elicit plant breeders' utility functions; however, numerous studies have attempted to characterize utility functions and risk preferences of agricultural producers (Lin et al. 1974; Wiens 1976; Moscardi and de Janvry 1977; Dillion and Scandizzo 1978; Binswanger 1980; Walker 1981). Most of these researchers first identified (either implicitly or explicitly) the functional form of the utility function and then estimated its curvature, which established the risk aversity of the producer. In this research the negative exponential utility function (1) was specified as the functional form, and the breeder's attitude toward yield stability was measured via the stability preference coefficient "a" ($a \geq 0$):

$$U(Y) = 1 - e^{-aY}. \quad (1)$$

This function (1) has several advantages. First, it adheres to the assumption of decreasing marginal utility of yield ($dU/dY > 0$; $d^2U/dY^2 < 0$). In addition, it results in a simple and intuitively meaningful index when yields are normally distributed (as will be assumed here). Finally, the negative exponential utility function has a single coefficient ("a"), which simply expresses the breeder's attitude toward yield stability; the larger the value of a, the more curved the utility function and the more importance the breeder places on stability of yield.

The coefficient "a" is directly related to the amount of yield required to make the breeder indifferent between a perfectly stable cultivar and an unstable cultivar. To understand this, assume the breeder has a concave utility function as shown in Fig. 1. Further assume that the yield of a perfectly stable cultivar (i.e., no yield variance) is always $\bar{Y}$, while yield Y of an unstable cultivar is a random variable, with realizations $\bar{Y}+e$ and $\bar{Y}-e$, each with probability ½. The utility the breeder gains from the perfectly stable cultivar is $U(\bar{Y})$, but the utility the breeder can expect to obtain from the unstable cultivar is its *expected* utility:

$$E(U(Y)) = ½[U(\bar{Y}+e) + U(\bar{Y}-e)]$$

as shown in Fig. 1.

Note in Fig. 1 that with a concave utility function, the expected utility the breeder obtains from the unstable cultivar is less than the utility he/she gains from the perfectly stable cultivar, even though both cultivars have the same mean yield. The difference between these two utilities is a measure of the 'cost' of instability when expressed in terms of the loss of expected utility. This 'cost' of instability, expressed in terms of yield, is $\bar{Y}-\hat{Y}$, as shown in Fig. 1. This difference, $\bar{Y}-\hat{Y}$, is called the risk premium, which represents the amount of yield from the perfectly stable cultivar that the breeder would have to give up, in order to make him/her indifferent to choosing between the perfectly stable cultivar and the unstable cultivar. In other words, the risk premium can be thought

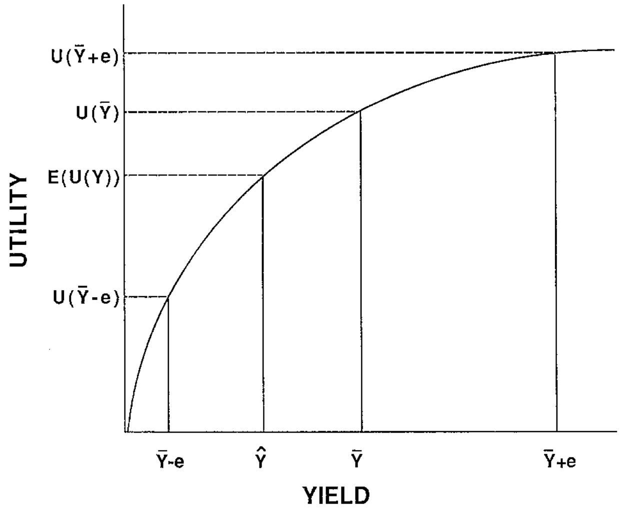

Fig. 1. Utility function of a risk-averse plant breeder

of as the 'cost' of yield instability expressed in terms of yield. The risk premium is approximately related to the stability preference coefficient ("a") as follows:

$$\bar{Y} - \hat{Y} = (a/2)\ V(Y), \quad (2)$$

where V(Y) is a variance measure of yield stability (Newbery and Stiglitz 1981). That is, the risk premium of an unstable cultivar is proportional to its yield stability. Half of the proportionality constant is the stability preference coefficient "a." When Y is normally distributed, Eq. (2) is exact.

Specification of probability distributions as 'predictions' of cultivar response

If each cultivar has constant yield in all environments, the 'best' cultivar is simply the one having the largest yield. However, yield varies from environment to environment, sometimes drastically. In addition, the breeder only uses a small sample of environments in which to evaluate cultivars targeted for all environments. To make selections in the presence of such yield uncertainty, the plant breeder must somehow predict yields for each cultivar in all environments. Application of EUM to selection requires that the breeder make these 'predictions' in terms of a probability distribution for each cultivar. Unfortunately, the shape or the true parameters (e.g., mean and variance) of the cultivars' yield distributions are rarely known. In this research, yield was assumed to be normally distributed, and sample estimates of the mean and variance from yield trials were used in place of the unknown true parameters. It should be noted that this assumption is consistent with assumptions that are made when selection is based on the more traditional criterion of estimates of means, measures of error, and tests of hypotheses made using analyses of variance.

Criterion for choice and selection of cultivars

To choose among several cultivars, the breeder must employ some sort of choice criterion. The most commonly used criterion is the largest expected (mean) yield. However, yield may not be a good reflection of the total value of a cultivar. A more general approach would be for the breeder to choose the cultivar that has the largest expected *utility,* i.e., the breeder uses EUM to choose among cultivars. Von Neumann and Morgenstern (1972) lent theoretical support to this mode of action by showing that if the breeder (or any decision maker) adheres to a few reasonable axioms, then he/she makes choices as if to maximize expected utility.

Selection indices based on EUM

The maximum expected utility criterion may be used to construct indices useful for selecting broadly adapted plant cultivars. To illustrate how such indices may be developed, let Y_{ij} represent the random yield for cultivar i ($i=1, \ldots, p$) in environment j ($j=1, \ldots, q$) with probability distribution $f(Y_{ij})$, and assume that the breeder values cultivars by using the utility function $U(Y_{ij})$. The 'value' the breeder may expect to obtain from the i^{th} cultivar is simply the expected value of the utility of yield:

$$E[U(Y_{ij})] = \int U(Y_{ij})\, f(Y_{ij})\, dY_{ij},$$

where the integration is over all possible yield values. If yield for the i^{th} cultivar Y_{ij}, is normally distributed with mean $E(Y_i)$ and variance $V(Y_i)$ and the breeder values cultivars using the utility function (1), then it can be shown (see Appendix) that the general form of an expected utility selection index is:

$$E(Y_i) - (a/2)\, V(Y_i). \tag{3}$$

The index value of the i^{th} cultivar is its expected yield minus its risk premium (2), which measures the cost of yield instability. The cultivar with the largest value of expression 3 is considered to be the 'best.'

If the general index form 3 is to be useful in aiding the plant breeder in selecting broadly adapted plant cultivars, $E(Y_i)$ must represent expected yield of cultivar i, and $V(Y_i)$ must represent the variance measure of stability of cultivar i based upon an acceptable stability model. Any one of a number of different stability models might be chosen, depending on the plant breeder's perspective. (There is still considerable debate on how to choose a stability model. See Lin et al. (1986) for a discussion of the advantages and disadvantages of the various approaches to stability.) Values of $E(Y_i)$ and $V(Y_i)$ based on stability models of Shukla (1972), Finlay and Wilkinson (1963), and Eberhart and Russell (1966) are given by Eskridge (1990). Once an acceptable stability model has been chosen, the appropriate estimates of $E(Y_i)$ and $V(Y_i)$, based on yield trial information, would be substituted into expression 3 and the cultivar with the largest value of expression 3 would be preferred.

Because of its wide acceptance and use among plant breeders, Eberhart and Russell's (1966) stability model will be used to illustrate the application of a specific EUM selection index. Using this model, it is possible to show that the i^{th} cultivar's adjusted yield has mean μ_i and variance $(\beta_i - 1)^2\, \sigma_{\bar{y}}^2\, (1 - 1/q) + \sigma_{\delta_i}^2$ (see Appendix). μ_i is the true mean yield for cultivar i, β_i is the i^{th} cultivar's slope coefficient obtained by regressing its mean yield on the mean yield of all cultivars for each environment, $\sigma_{\bar{y}}^2$ is the variance of environment means, and $\sigma_{\delta_i}^2$ is the i^{th} cultivar's variance of the deviations about regression. Using the breeding trial data, μ_i is estimated by the sample mean ($\bar{Y}_{i\cdot}$), β_i and $\sigma_{\delta_i}^2$ are estimated with b_i and $S_{\delta_i}^2$ as in Eberhart and Russell (1966), and $\sigma_{\bar{y}}^2$ is estimated with the sample variance of environmental means. These quantities are substituted for $E(Y_i)$ and $V(Y_i)$ in expression 3, and the resulting specific form of the expected utility selection index is given as follows:

$$\bar{Y}_{i\cdot} - (a/2)\, [(b_i - 1)^2\, S_{\bar{y}}^2\, (1 - 1/q) + S_{\delta}^2]\,. \tag{4}$$

Assuming that Eberhart and Russell's (1966) approach is an acceptable stability model and that the plant breeder's preferences can be approximated with the utility function given in Eq. 1 where his/her stability preference coefficient ("a") is known, the cultivar with the largest value of the index 4 is considered to be the 'best.'

Index 4 could be viewed as a special form of Williams' (1962) base index; however, it is more straightforward to consider index 4 to be an EUM index developed from a base index, with one trait (yield) with a weight of one. Note that index 4 does not consider genotypic variances and covariances as does the traditional Smith (1936) index or any of its modifications. In this instance, the breeder is choosing among a set of cultivars, largely unrelated, for final release. Selection is not being made from among many genotypes taken from some reference population, either actual or arbitrary. Consequently, in this index the assumption is made that genetic effects are fixed.

Regional uniform soybean trials

USDA Regional Uniform Soybean Test data were used to illustrate how EUM might be used to aid the plant breeder in selecting broadly adapted cultivars. During the years 1954–1956, test varieties in several maturity groups were held constant with the purpose of improving estimates of genetic parameters (Schutz and Bernard 1967). Maturity group zero, which comprises varieties adapted to conditions in the North American upper mid-west was

Table 1. Soybean varieties from maturity group 0, 1954–1956 regional uniform soybean tests

No.	Variety	Source	Pedigree
1	W0S-3386	Wis. AES/ Reg. SB. Lab.	Lincoln × Flambeau
2	Chippewa	Ill. AES/ Reg. SB. Lab.	Lincoln × (Linc. × Rich.)
3	Hardrome	Dominion Exp. Farm	Mandarin × (Man. × A.K.)
4	W0-3147	Wis. AES/ Reg. SB. Lab.	Mukden × Flambeau
5	W92-2703	Wis. AES/ Reg. SB. Lab.	Lincoln × Flambeau
6	Capital	Central Exp. Farm, Ottawa	Strain 171 × A.K.
7	Comet	Central Exp. Farm, Ottawa	Pagoda × Mandarin
8	Renville	Minn. AES/ Reg. SB. Lab.	Lincoln × (Linc. × Rich.)
9	W0S-3257	Wis. AES/ Reg. SB. Lab.	Mukden × Flambeau
10	W0S-3180	Wis. AES/ Reg. SB. Lab.	Mukden × Flambeau
11	Mandarin	Central Exp. Farm, Ottawa	Mandarin
12	W0S-3138	Wis. AES/ Reg. SB. Lab.	Hawkeye × Flambeau
13	Norchief	Wis. AES/ Reg. SB. Lab.	Hawkeye × Flambeau
14	Flambeau	Wis. AES	Intr. from USSR

Table 2. Locations used in the 1954–1956 regional uniform soybean tests, maturity group 0

Location no.	City	State/Province
1	Ottawa	Ontario
2	Guelph	Ontario
3	Spooner	Wisconsin
4	Durand	Wisconsin
5	Morris	Minnesota
6	St. Paul	Minnesota
7	Casselton	North Dakota
8	Fargo	North Dakota

used. The layout was a randomized complete block design with four blocks and 14 varieties, evaluated at each of eight locations over the 3-year period, resulting in 24 environments. Varieties were assumed to be fixed, as opposed to a random sample of varieties. Table 1 lists the soybean varieties, their sources and origins, while Table 2 lists the locations. Measurements were available on yield, oil, maturity, lodging, height, seed quality, and seed weight.

Genotypic rankings and stability preference

For simplicity, yield was assumed to be the only trait of interest. To apply the above selection index, $\bar{Y}_{i\cdot}$, b_i, and

Table 3. Mean yields, regression coefficient (b_i), and mean squares deviation about regression ($S^2_{\delta i}$) for each variety and variance of environmental means ($S^2_{\bar{y}}$) for varieties in maturity group 0 regional uniform soybean test, 1954–1956

Variety	Mean[a]	b_i	$S^2_{\delta i}$
1	1.935	1.103	0.015
2	1.911	1.384	0.053
3	1.888	1.015	0.027
4	1.878	1.000	0.024
5	1.801	0.860	0.016
6	1.857	1.057	0.034
7	1.797	0.872	0.016
8	1.800	1.194	0.041
9	1.772	1.015	0.016
10	1.779	0.838	0.023
11	1.799	1.030	0.026
12	1.725	0.929	0.017
13	1.735	0.930	0.014
14	1.616	0.772	0.061
	$S^2_{\bar{y}}=0.253$		

[a] Milligrams per hectare

$S^2_{\delta_i}$ were used to estimate μ_i, β_i, and $\sigma^2_{\delta_i}$ for each variety. These quantities for the regional uniform soybean tests are given in Table 3, where each location-year combination is treated as an environment. Yields are considered normally distributed, since residuals from a model including variety and environment as classification variables are concluded to be normal, using the Kolmogorov-Smirinov test for normality ($P>0.01$).

The order, or ranking, of the soybean varieties depends on the functional form of the breeder's utility function, its curvature, and each variety's yield distribution. The breeder's utility function was assumed to be represented by the negative exponential function 1. To use this utility function with EUM, it is necessary to elicit the breeder's stability preference coefficient, "a", which defines the curvature of the function 1. The breeder's "a" value may be directly elicited from the plant breeder using any of several different approaches (Anderson et al. 1977). Assuming the breeder chooses the Eberhart and Russell (1966) approach as the preferred stability model, and using the breeder's elicited "a" value, the cultivar with the largest value of the index 4 would be preferred.

To illustrate how different levels of importance placed on yield stability affect cultivar ordering, different stability preference values ("a") were used in index 4 to correspond with severe, intermediate, moderate, slight, and no levels of concern about yield stability (severe: $a=4.1505$; intermediate: $a=0.9684$; moderate: $a=0.4427$; slight: $a=0.1660$; none: $a=0$). These particular values of "a" correspond to the range of income risk attitudes of agricultural producers found in experiments conducted by (Binswanger 1980). Other researchers studying agricul-

Table 4. Variety index values and rankings (in parentheses) at five different stability preference levels

Variety	Level of concern about yield stability				
	None (0.00)[a]	Slight (0.1660)	Moderate (0.4427)	Intermediate (0.9684)	Severe (4.1505)
1	1.935 (1)	1.933 (1)	1.931 (1)	1.926 (1)	1.897 (1)
2	1.911 (2)	1.904 (2)	1.892 (2)	1.868 (3)	1.727 (9)
3	1.888 (3)	1.885 (3)	1.882 (3)	1.875 (2)	1.832 (2)
4	1.878 (4)	1.876 (4)	1.873 (4)	1.867 (4)	1.829 (3)
5	1.801 (6)	1.799 (6)	1.796 (6)	1.791 (6)	1.757 (5)
6	1.857 (5)	1.855 (5)	1.850 (5)	1.840 (5)	1.785 (4)
7	1.797 (9)	1.795 (9)	1.792 (8)	1.787 (7)	1.754 (6)
8	1.800 (7)	1.796 (8)	1.789 (9)	1.776 (9)	1.697 (12)
9	1.772 (11)	1.770 (11)	1.768 (11)	1.764 (11)	1.738 (8)
10	1.779 (10)	1.777 (10)	1.773 (10)	1.765 (10)	1.718 (10)
11	1.799 (8)	1.797 (7)	1.793 (7)	1.787 (8)	1.746 (7)
12	1.725 (13)	1.723 (13)	1.721 (13)	1.716 (13)	1.687 (13)
13	1.735 (12)	1.733 (12)	1.731 (12)	1.727 (12)	1.702 (11)
14	1.616 (14)	1.610 (14)	1.600 (14)	1.580 (14)	1.463 (14)

[a] Values of "a" associated with each level of concern about yield stability

Table 5. Kendall rank correlations between variety rankings, using different levels of concern about yield stability

	Level of concern about yield stability				
	None (0.00)[a]	Slight (0.1660)	Moderate (0.4427)	Intermediate (0.9684)	Severe (4.1505)
None	1.000	0.978	0.956	0.912	0.692
Slight		1.000	0.978	0.934	0.714
Moderate			1.000	0.956	0.736
Intermediate				1.000	0.780
Severe					1.000

[a] Values of "a" associated with each level of concern about yield stability

tural producers' perceptions of income risk have observed similar ranges of attitudes (Dillon and Scandizzo 1978; Walker 1981). It is assumed that this range of risk attitudes corresponds to the range of values for preference of stability exhibited by different plant breeders. Binswanger (1980) expresses attitudes towards risk in terms of partial risk aversion, "P", which is a relative measure of risk aversion (Newbery and Stiglitz 1981). The "a" values for the different levels of concern about yield stability were obtained from the relationship $P = \bar{Y}.. \, a$ where $\bar{Y}..$ is the mean yield from the soybean trial, i.e., $\bar{Y}.. = 1.807$ mg/ha.

Table 4 gives the index values and ranks of the soybean varieties for the five levels of stability preference. Parenthetical numbers after each value represent each variety's rank. Different rankings illustrate how different stability preferences affected the relative desirability of the different varieties. If the plant breeder is not concerned about yield stability, then his/her stability preference coefficient, "a," is zero in index 4 and selections are based on the mean yield alone. However, if the breeder is concerned about stability, then "a" > 0 and stability affects the relative desirability of a variety, with larger values of "a" being associated with more concern with stability. Table 5 gives rank correlations between variety rankings associated with the five different levels of concern about stability: the greater the difference in concern about stability, the more two rankings will differ and the smaller is the rank correlation coefficient.

A few of the major reversals in rankings due to different levels of concern about stability are worth noting. Chippewa, variety 2, was second best in terms of mean yield. However, as stability became more important, Chippewa became less valuable, eventually dropping to ninth at the severe level of concern about stability. This reversal was due to Chippewa's large stability variance, which appeared to be caused by high yields in better environments and poor yields in marginal environments. A similar rank reversal, although not as extreme as with Chippewa, was also observed with Renville, variety 8.

Discussion

Models based on EUM were first presented in the statistics and economics literature and have been used extensively in business applications to model how people make choices. The EUM approach to decision making implicitly assumes that the decision maker evaluates each alternative with his/her own utility function and prefers the alternative with the largest average (or expected) utility value. The EUM approach used in this research was based on the assumptions that the decision maker's (plant breeder) utility function was known and that the payoff variable (yield) was continuous with a known probability distribution. This approach is analogous to the analysis of the utility of continuous payoff distributions, as described in Bunn (1984), and is more general than the elementary payoff matrix approach to EUM, as discussed in most elementary applications.

EUM provides an approach to decision making that can be used to develop selection indices useful for selecting broadly adapted plant cultivars. Using utility concepts from decision theory and a stability model the breeder finds acceptable, a selection index can be developed that explicitly quantifies how the plant breeder weighs the importance of yield relative to stability. A specific index 4 is developed, with stability preferences being reflected by the magnitude of the stability preference coefficient "a." This index could be used in advanced testing, where elite, but mostly unrelated, materials are being tested for possible release.

The regional soybean testing data demonstrated how this specific EUM index 4 could be used to select broadly adapted plant cultivars. Comparison of the variety rankings indicated that different levels of concern about stability produced different orderings. The rankings of some varieties were quite sensitive to the breeder's level of concern about instability. Such major reversals were possible since variety ordering can be quite sensitive to the breeder's attitude towards cultivar instability.

The choice of a particular definition of stability may also affect cultivar rankings. In this study, Eberhart and Russell's (1966) stability model was used, since this approach has been widely used by practicing plant breeders, even though it has been criticized (Lin et al. 1986). Other stability approaches, however, could be incorporated into the general EUM index 3. Indices based on Shukla's (1972) or Finlay and Wilkinson's (1963) stability approaches could be easily developed by estimating $E(Y_i)$ and $V(Y_i)$, as shown in Eskridge (1990), and substituting these values in expression 3. More complicated models that account for both sites and years could also be developed after obtaining appropriate estimates of $E(Y_i)$ and $V(Y_i)$. Although further investigation is necessary, EUM indices based on approaches such as the AMMI model (Gauch 1988) and the shifted multiplicative model (Seyedsadr and Cornelius 1989) should also be possible.

From a more general perspective, there are several advantages in using EUM in the development of selection indices. First, selection of the 'best' cultivar involves choosing among cultivars with uncertain performance in many environments. EUM, and decision theory in general, was developed to provide a framework for such choices (Bunn 1984; Anderson et al. 1977). In addition, EUM is very general in that the paradigm can accommodate most methods of selection and nearly any phenotypic worth index (linear or nonlinear indices of economic worth, merit, profit, etc). In this research, EUM was used to develop a selection index useful for selecting broadly adapted cultivars in the final stages of a breeding program, when yield is the only trait being considered. However, other applications are also possible. For example, EUM could easily be applied to population improvement using a multi trait Smith-Hazel index. The breeder would generate his/her index of total phenotypic value as per Smith (1936), using economic weights and genetic variances and covariances of the traits. Then the index's phenotypic value would be used in place of yield, along with the measure of the stability of the index value, to obtain an EUM index.

Another advantage of EUM is that it allows for incorporation of different stability models and different stability preferences. The ability to incorporate different perspectives on stability in a selection scheme allows cultivars to be selected that are more likely to fit the breeder's definition of 'best.' Although not discussed, a further advantage of using EUM in selection is the ability of the model to explicitly incorporate the breeder's prior knowledge of the cultivars in the selection rules. Anderson et al. (1977) discuss how prior knowledge can be incorporated into EUM, and Gianola and Fernando (1986) illustrate how prior knowledge may be used in breeding and selection.

There are several disadvantages in applying EUM to selection. First, the EUM indices aid the breeder in selecting cultivars in the presence of genotype × environment interaction; however, the breeder should not use these indices in lieu of understanding the biological nature of genotype × environment interactions. Second, the breeder's utility function must be specified. This involves both specifying the breeder's utility function and estimating his/her stability preference coefficient(s). Methods exist for eliciting a plant breeder's stability preference coefficient(s), assuming that the form of the utility function is known (Anderson et al. 1977), but the form and properties of utility functions are still being debated (Kahneman and Tversky 1982; Schoemaker 1982). In addition, application of EUM to selection requires that the breeder specify a probability distribution for each cultivar, yet this practice is not uncommon. Plant breeders have long been specifying distributions when they make single-trait selection based on analysis of variance. Finally, assuming that the form of the probability distributions is known, the parameters of these distributions could be very poorly estimated, especially if the estimates are based on a small number of environments. Nevertheless, application of EUM to selection can be a useful tool to the plant breeder attempting to select broadly adapted cultivars since: (1) the importance of yield to stability is explicitly stated through the specified utility function, (2) utility-based selection indices are more likely to identify superior varieties when stability is a major consideration, and (3) EUM may be used with any type of worth index and nearly any breeding plan.

Acknowledgements. The authors would like to thank Dr. J. R. Wilcox and the various cooperators for providing the data used in this investigation, and the reviewers for helpful comments and suggestions.

Appendix

Derivation of the expected utility selection index

If $U(Y)=1-e^{-aY}$ then, as shown in Freund (1955):

$$\mathrm{E}(U(Y))=E(1-e^{-aY})=1-\mathrm{E}(e^{-aY})=1-M(-aY),$$

where $M(.)$ is the moment-generating function of the probability distribution of Y. If Y is assumed to be normally distributed,

$$M(-aY)=e^{[-a\,\mathrm{E}(Y)+(a^2\,\mathrm{V}(Y)/2)]}.$$

Therefore, choosing the cultivar to maximize $\mathrm{E}(U(Y))$ is the same as choosing the cultivar that has the largest value of

$E(Y)-(a/2)V(Y)$. However, it is being assumed that the true parameters equal the sample estimates, thus the value of the selection index for cultivar i is $\bar{Y}-(a/2)S_{\bar{Y}}^2$.

Stability model based partially on Eberhart and Russell (1966)

This approach characterizes the desirability of a cultivar using three parameters: its mean yield (μ_i), regression coefficient (β_i), and the mean square deviations about regression($\sigma_{\delta i}^2$). Following Eberhart and Russell (1966), define the following model:

$$Y_{ij}=\mu_i+\beta_i(\bar{Y}_{.j}-\bar{Y}_{..})+\delta_{ij},$$

where μ_i is the true mean for cultivar i, β_i is the regression coefficient for cultivar i, $\bar{Y}_{.j}$ is the marginal mean of environment j, which is assumed to be a random variable with true mean $\mu_{\bar{y}}$ having variance $\sigma_{\bar{y}}^2$. δ_{ij} is a normally distributed error, which is independent of $\bar{Y}_{.j}$, with mean 0 and variance $\sigma_{\delta i}^2$. By adding and subtracting the product of the mean slope and the 'environment index,' $\beta(\bar{Y}_{.j}-\bar{Y}_{..})$, to the right side of this equation, we obtain:

$$Y_{ij}=\mu_i+\beta(\bar{Y}_{.j}-\bar{Y}_{..})+(\beta_i-\beta)(\bar{Y}_{.j}-\bar{Y}_{..})+\delta_{ij}.$$

However, $\beta(\bar{Y}_{.j}-\bar{Y}_{..})$ contains no information about the i^{th} cultivar, since it represents how the average cultivar responds to the environment index. Therefore, subtracting $\beta(\bar{Y}_{.j}-\bar{Y}_{..})$ from both sides of the equation gives the following 'adjusted' yield (YA_{ij}), which contains all the information about the i^{th} cultivar relevant to the Eberhart and Russell (1966) approach:

$$YA_{ij}=\mu_i+(\beta_i-\beta)(\bar{Y}_{.j}-\bar{Y}_{..})+\delta_{ij}.$$

Using rules of expectation and variance and since the mean slope (β) is always 1, the i^{th} cultivar's population mean and variance of the adjusted yield over all environments can be shown to be μ_i and $(\beta_i-1)^2\,\sigma_{\bar{y}}^2(1-1/q)+\sigma_{\delta i}^2$, respectively.

References

Allaire FR, Thraen CS (1985) Prospects for genetic improvement in the economic efficiency of dairy cattle. J Dairy Sci 68:3110–3123

Anderson JR, Dillon JL, Hardaker JB (1977) Agricultural decision analysis. Iowa State University Press, Ames/IA

Barah BC, Binswanger HP, Rana BS, Rao NGP (1981) The use of risk aversion in plant breeding: concept and application. Euphytica 30:451–458

Binswanger HP (1980) Attitudes toward risk: theoretical implications of an experiment in rural India. Econ J 91:867–891

Bunn DW (1984) Applied decision analysis. McGraw-Hill, New York

Dillon JL, Scandizzo PL (1978) Risk attitudes of subsistence farmers in northeast Brazil: a sampling approach. Am J Agric Econ 60:425–435

Eberhart SA, Russell WA (1966) Stability parameters for comparing varieties. Crop Sci 6:36–40

Eskridge KM (1990) Selection of stable cultivars using a safety first rule. Crop Sci 30:369–374

Finlay KW, Wilkinson GW (1963) The analysis of adaptation in a plant breeding program. Aust J Agric Res 14:742–754

Freund RJ (1955) Introduction of risk into a programming model. Econometrica 24:253–263

Gauch HG (1988) Model selection and validation for yield trials with interaction. Biometrics 44:705–716

Gianola D, Fernando RL (1986) Bayesian methods in animal breeding theory. J Anim Sci 63:217–244

Kahneman D, Tversky A (1982) The psychology of preferences. Sci Am 246:160–173

Lin W, Dean GW, Moore CV (1974) An empirical test of utility versus profit maximization in agricultural production. Am J Agric Econ 56:497–508

Lin CS, Binns MR, Lefkovitch LP (1986) Stability analysis: Where do we stand? Crop Sci 26:894–899

Moscardi ER, Janvry A de (1977) Attitudes toward risk among peasants: an econometric approach. Am J Agric Econ 59:710–716

Neumann J von, Morgenstern O (1972) Theory of games and economic behavior. Princeton University Press, Princeton/NJ

Newbery DMG, Stiglitz JE (1981) The theory of commodity price stabilization: a study in the economics of risk. Clarendon Press, Oxford

Schneeberger M, Freeman AE, Boehlje MD (1982) Application of portfolio and utility theories for dairy breeding. J Dairy Sci 65:404–9

Schoemaker PJH (1982) The expected utility model: its variants, purposes, evidence, and limitations. J Econ Lit 20:529–563

Schutz WM, Bernard RL (1967) Genotype × environment interactions in the regional testing of soybean strains. Crop Sci 7:125–130

Seyedsadr M, Cornelius PL (1989) Estimation of parameters of the shifted multiplicative model for a two-way table. Stat Dep Tech Rep 278, University of Kentucky, Lexington/KY

Shukla GK (1972) Some statistical aspects of partitioning genotype-environment components of variance. Heredity 29:237–245

Smith HF (1936) A discriminant function for plant selection. Ann Eugenics 7:232–248

Smith SP, Allaire FR (1985) Efficient selection rules to increase nonlinear merit: an application to mate selection. Genet Sel Evol 17:387–406

Smith SP, Hammond K (1986) The ramification of portfolio and utility theories for dairy breeding programs. Proc 3rd World Congr Genet Appl Livestock Prod, August 1986, Lincoln NE

Walker TS (1981) Risk and adoption of hybrid maize in El Salvador. Food Res Inst Stud 18:59–88

Wiens TB (1976) Peasant risk aversion and allocative behavior: a quadratic programming experiment. Am J Agric Econ 58:629–635

Williams JS (1962) The evaluation of a selection index. Biometrics 18:375–393