Mem. Fac. Agr. Kagoshima Univ., **21**, 67–75, 1985

Evaluation of the Efficiencies of Selection Indices with Erroneous Parameters

Muneharu SATO, Anton J. BOCKHOLT* and Teiji KOKUBU
(*Laboratory of Plant Breeding*)
Received for Publication September 10, 1984

Introduction

The concept of an optimum selection criterion was proposed by Smith[17] and the selection index based on it was established by Hazel[6]. To meet the diverse needs of breeders, various types of indices have been developed since then. Kempthorne and Nordskog[7], for example, introduced a restricted selection index. Later it was refined by Cunningham *et al.*[4] To solve the nonlinearity of aggregated genotype, a nonlinear selection index such as a quadratic index was also originated.[21] Pesek and Baker[11] solved the ambiguity of economic weights by introducing an economic weight free selection index.

Selection indices have been being evaluated as the possible alternative of relatively inefficient single trait selection methods by the field experiments. For example, Elgin, *et al.*[5] compared four methods of multiple selection, tandem, independent culling levels, Williams' base index, and the Smith-Hazel index for five characters in alfalfa (*Medicago sativa* L.). And they found that the base index and the Smith-Hazel index were most effective. Five different types of indices were compared by Subandi *et al.*[19] in their efficiencies for yield improvement. Two types of Smith-Hazel indices with different economic weights were most effective, depending on the harvesting methods.

Generally speaking, however, selection indices have not been applied for plant breeding extensively. Comstock[1] stated the probable reasons for the unpopularity of the selection index: 1. costly data collection, 2. assumption of single stage selection procedure, and 3. uncertain effects of sampling errors in genetic parmeter estimates.

Currently the cost related problems are still difficult to solve, though automated data collection systems may partially relieve the difficulties in the near future. But in some situations, the selection index may be cheaper. For example, the selection for sucrose yield of sugar cane by the selection index with alternative characters was found to be sufficiently effective and more economical than the direct single selection method[8]. To solve the single stage restriction, a multistage selection index was proposed[3].

The ambiguity of effects of inaccurate parameters, on the other hand, has not been clarified yet, though Pease[10] indicated that the likely error in parameters obtainable at present for pigs in England could cause a loss in a rate of progress of approximately five percent and certainly less than 10 percent. Therefore, the objective of this research is to investigate the effects of erroneous parameters of selection indices on their efficiencies.

* Department of Soil and Crop Sciences, Texas A & M Univ., College Station, Texas, U. S. A.

Materials and Methods

First the parameters for the reference index, or the error free index to be compared with erroneous indices, were obtained from published data[9,12,13,14] on maize (*Zea mays* L.). The values chosen were the values which appeared most frequently or the values believed to be appropriate. As virtually no studies on the economic weights were found, in this study the weights, 0, 0, 0, 10, 0, 0, and 1 were used for plant height, ear height, husk length, ear number/plant, ear length, ear diameter, and yield, respectively. Since the profit accounted for the yield was considered here, the weight for the ear number, which is discrete integer number, was determined to be 10. This value, however, may differ from population to population. It also depends on the definition of the profit. Therefore, the values were only intended to be used to give the starting point to construct the reference index.

The parameters used to construct the reference index in this study are shown in Table 1.

Table 1. Paramaters used to construct the reference Smith-Hazel selection index

Character Parameter	Plant height (cm) X1	Ear height (cm) X2	Husk length (cm) X3	Ears/ plant X4	Ear length (cm) X5	Ear diameter (cm) X6	Yield (kg/plant) X7
Heritability	0.07	0.05	0.06	0.02	0.15	0.16	0.15
Phenotypic variance	186.45	106.46	2.80	0.0525	0.8529	0.0329	0.00144
Economic weight	0.	0.	0.	10.	0.	0.	1.
Correlation*							
X1	1.	0.749	0.157	0.188	0.054	0.025	0.216
X2	0.840	1.	0.072	0.213	−0.058	−0.007	0.266
X3	0.299	0.144	1.	0.013	−0.254	−0.221	−0.275
X4	0.300	0.205	0.148	1.	0.021	−0.126	0.691
X5	0.049	0.028	−0.037	−0.030	1.	0.313	0.393
X6	−0.066	−0.069	−0.412	−0.075	0.184	1.	0.298
X7	0.381	0.478	−0.112	0.819	0.188	0.174	1.

* The upper off-diagonal elements and the lower off-diagonal elements of the correlation matrix represent phenotypic and genetic correlation coefficients of two characters, respectively.

The coefficient vector $\boldsymbol{b}$ and the gain $\boldsymbol{H}$ of the reference index were derived by the following formulae.

$$\boldsymbol{b}=\boldsymbol{P}^{-1}\boldsymbol{G}\boldsymbol{a},$$

$$\boldsymbol{H}=i\,(\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}),$$

where $\boldsymbol{P}$ designates the phenotypic variance and covariance matrix, $\boldsymbol{G}$ represents the genotypic variance and covariance matrix, $\boldsymbol{a}$ is the economic weight vector, and i is the selection intensity. Superscripts $^{-1}$ and $'$ show the inverted matrix and the transposed vector, respectively.

The efficiencies of the reduced index models were also evaluated by the method proposed by Cunningham[2]:

The reduction of the efficiency if the variate is omitted

$$=100[1-\{(b'Pb-b_i^2/w_{ii})/b'Pb\}]^{1/2},$$

where b_i is the i-th coefficient in the vector $\boldsymbol{b}$ or the coefficient of i-th trait which is removed from the model, and w_{ij} is the diagonal element of the inverted phenotypic variance and covariance matrix (P^{-1}).

The relative efficiency of the index with erroneous parameters were examined by taking the ratio of the aggregated genetic gain, expected from the selection based on the index obtained from the incorrect parameters and that based on the reference index, which will eventually be

$$\hat{b}Ga/(\hat{b}'P\hat{b}b'Pb)^{1/2},$$

where $\hat{\boldsymbol{b}}$ is the estimated coefficient vector of a faulty index.

Results and Discussion

1. Evaluation of Smith-Hazel selection index

The efficiency of the reference selection index used as the standard correct index is contrasted with that of a single trait selection method of each of seven characters in Table 2. Though the best single trait selection method is to select for yield, it is only 65.8 percent as effective as the selection index method as shown in the column of relative efficiency. The selection on ear length or ear diameter actually reduced the aggregated genetic gain.

Table 2. The efficiency of a Smith-Hazel selection index

	Selection	Index	Selection	Single character selection	
Plant characters	Index cofficients	Expected gain	Cost of reduction	Relative*2 efficiency	Expected gain
Plant height (cm)	0.003	$0.905i$*1	1.5%	16.5%	$0.956i$
Ear height (cm)	-0.007	$0.691i$	4.0	9.9	$0.516i$
Husk length (cm)	0.034	$0.033i$	4.8	6.9	$0.100i$
Ears/plant	-0.736	$0.016i$	19.7	28.8	$0.005i$
Ear length (cm)	-0.094	$-0.037i$	9.9	negative	$0.139i$
Ear diameter (cm)	-0.464	$-0.010i$	9.5	negative	$0.029i$
Yield/plant (kg/plant)	8.157	$0.007i$	67.9	65.8	$0.006i$

Aggregated genetic gain $0.165i$

*1 'i' designates the selection intensity (i=selection differential/standard deviation).

*2 relative efficiency$=100\times$aggregated genetic gain by the single character selection/0.1651.

Interestingly enough, in ear height, ear number and yield, it was found that the selection index method was expected to confer greater response than the single trait selection method for these characters. The relatively high genetic correlations between them would be responsible for this phenomenon. The result suggests that the selection index, which allows all characters to change in optimum mode, could intensify the response of important characters more than the single trait selection on them.

The cost of reduction in this table indicates how much the efficiency of the index decreases if the particular character is ignored from the observation. For example, if the ear number is removed from the model to reduce the model, the efficiency would drop to 80.3 percent of the original seven-character model. The results show that plant height is the best candidate to be removed from the observation, if necessary. It may not be recommended, on the other hand, to omit measurement of yield so long as breeders desire the sufficient aggregated genetic progress.

2. Errors in heritabilities

As the square roots of the elements of heritability vector are multiplied with each other to determine genetic variance-covariance matrix, no change is expected in the efficiency of the index if every element is biased with equal proportion. As shown in Fig. 1, both single and compound errors in heritabilities seem to have insignificant effect on the effectiveness of the index. Curiously enough, the compound errors in ear number/plant and yield actually compensated to each other and reduced the loss of the efficiency of the index in this particular model.

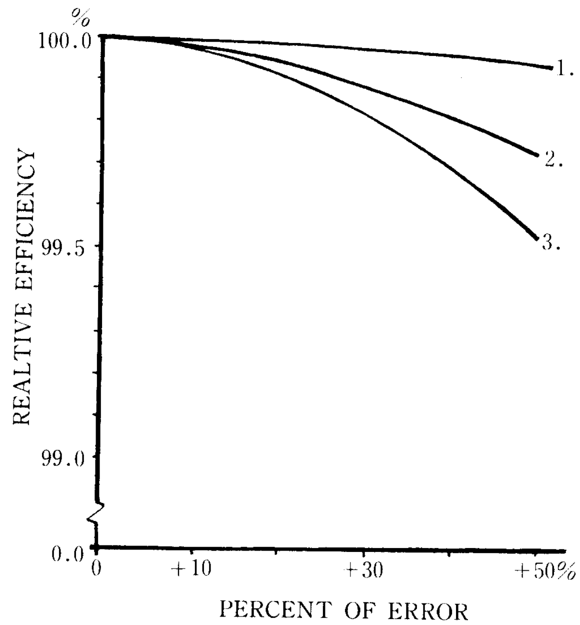

Fig. 1. The relative efficiency when heritabilities are affected by errors. The lines 1, 2 and 3 represent the changes in the efficiency when errors are in heritabilities of ear and yield, of yield and of ear number, respectively.

In both plants and animals, the degrees of dominance and epistasis which are considered to be major causes of biases in heritability estimates were reported to be less than 10 percent in most cases[12,14,15,16,18]. Therefore, if the experiment is carefully executed and the inbreeding effect is corrected, the heritability estimates obtained could be used to construct the selection index without any significant reduction in its effectiveness.

3. Errors in genetic correlations

The changes in the efficiency of the index when the genetic correlations are subject to errors are illustrated in Fig. 2. The underestimated parameter reduced the efficiency more than the overestimated ones, though less than 10 percent of the loss was estimated in all cases. The combined errors in both heritability estimates of ear number and of yield, and their genetic correlation estimates are also presented in this figure. As expected from the results of the effects of errors in heritabilities, the effectiveness reduced just like the case in which the error was in the correlation

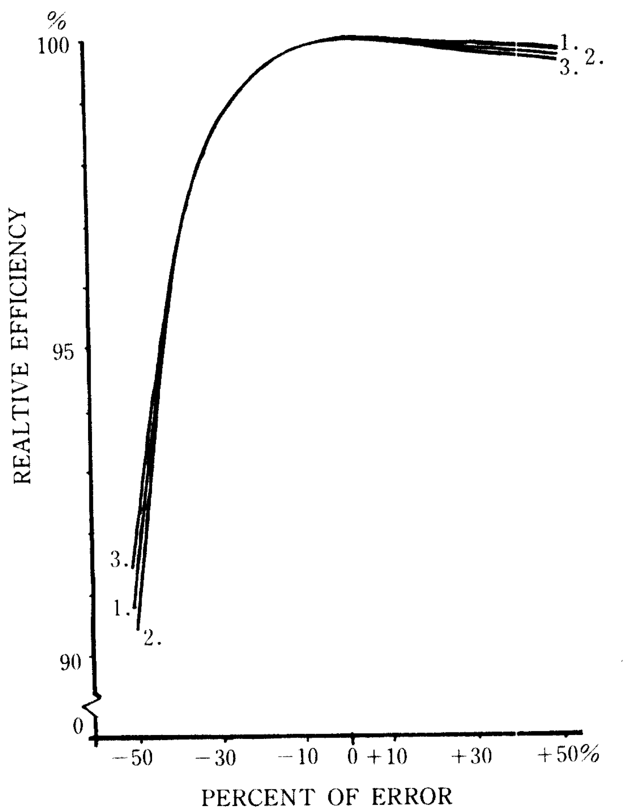

Fig. 2. The relative efficiency when genetic correlations and heritabilities are affected by errors. The lines 1, 2 and 3 represent the changes in the efficiency when errors are in genetic correlation between ear number and yield, in all off-diagonal elements of genetic correlation matrix and in genetic correlation between ear number and yield, heritability of ear number and heritability of yield, respectively.

alone. The effects of the underestimated correlations, however, were slightly alleviated by the positively deviated heritability estimates.

Theoretically, the i-th row of the j-th column element of the genetic variance-covariance matrix (g_{ij}) is defined as $g_{ij}=r_{ij}h_i h_j \sigma_i \sigma_j$,

where r_{ij}, h_i, h_j, σ_i, and σ_j are the genetic correlation between the i-th character and the j-th character, the square root of heritability of the i-th character and that of the j-th character, the standard deviation of the i-th character and that of the j-th character, respectively.

Hence, underestimated r_{ij} could be compensated by the inflated heritability estimates to obtain the g_{ij} estimate with smaller bias.

4. Errors in phenotypic variance

The proportional errors in all elements have no effect on the relative efficiency of the index by the same reason as those in heritabilities. Fig. 3 shows the changes when the faulty phenotypic variances are used. Both single element and compound errors provided the efficiency of over 90 percent of the true index. The compound errors did not affect the efficiency of the index any more than a single error.

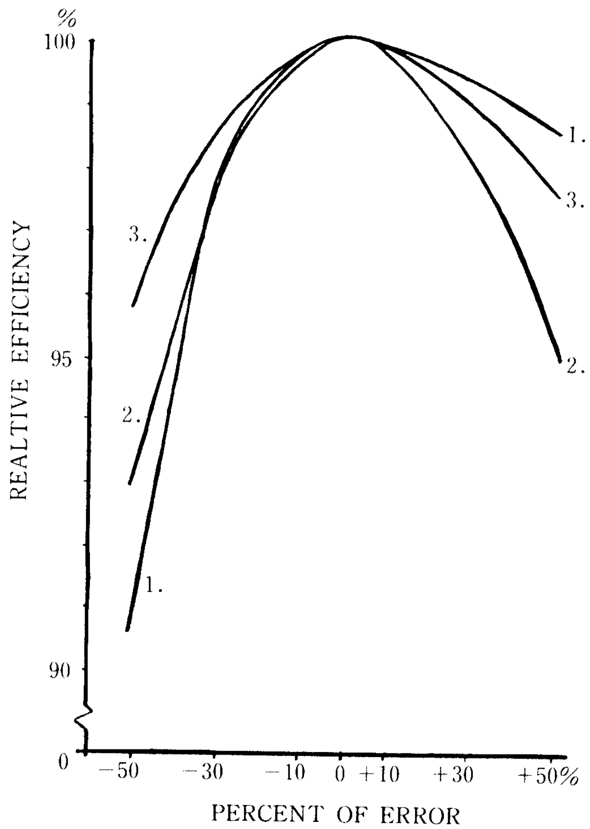

Fig. 3. The relative efficiency when phenotypic variances are affected by errors. The lines 1, 2 and 3 represent the changes in the efficiency when errors are in phenotypic variances of ear numbers, yield, and of ear number and yield, respectively.

5. Errors in phenotypic correlations

The relative efficiencies of the indices obtained by the faulty phenotypic correlations are indicated in lines 1 and 2 in Fig. 4. When some of the phenotypic correlations were over-estimated by 23.5 percent or more, the indices actually hindered the total genetic gains. For example, if the true correlation between ear number and yield, 0.691, was erroneously estimated to be over 0.853, no progress was observed. This phenomenon suggests that the misinterpretation of the intermediate relationship between two characters as the highly correlated association would severely reduce the efficiency of the index. The underestimated parameters, however, have less impact on the efficiency of the index than the over-estimated ones. Even if the parameters were underestimated by 50 percent, the 75 percent of the efficiency of the optimal index was maintained.

6. Errors in economic weights

If all economic weights are subject to errors in the same proportion, there will be no effect in the efficiency as in the cases of heritabilities and phenotypic variances. As demonstrated in Fig. 5, both single element errors and compound errors in the weights of yield and ear number produced no significant effect on the efficiencies.

Maximum loss was observed when both economic weights of ear number and yield were underestimated by 50 percent but loss in relative efficiency was still less than 0.03 percent. As Vandepitte and Hazel[20] obtained the similar results from the study on the selection index for pigs, the results of this research may not be unique.

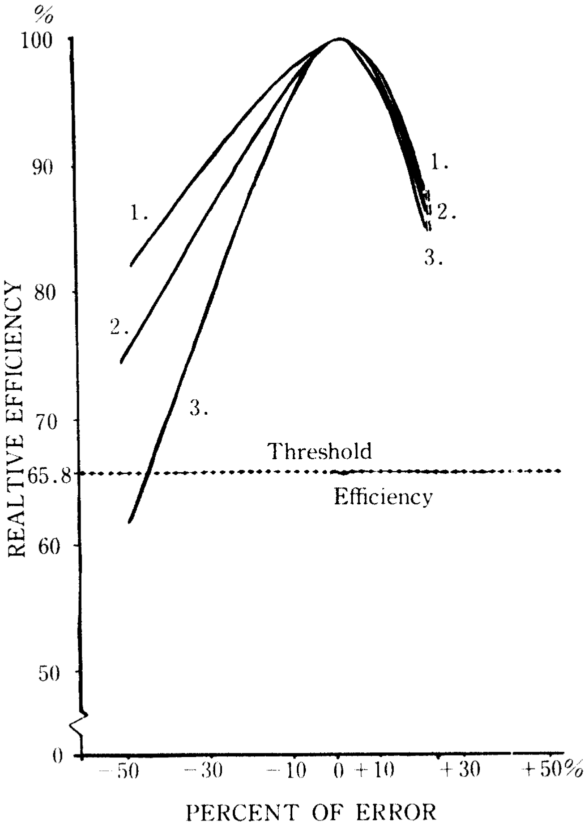

Fig. 4. The relative efficiency when phenotypic correlations and heritabilities are affected by errors. The lines 1, 2 and 3 represent the changes in the efficiency when errors are in phenotypic correlation between ear number and yield, in all diagonal elements of phenotypic correlation matrix, and in heritability of ear number, that of yield and all off-diagonal elements of phenotypic correlation matrix, respectively.

Although economic weights may become ambiguous when the prices of grains in future are unpredictable, they are actually defined in the terms relative to other traits. Therefore, unless farming methods and market structures change radically, the relative proportions among economic weights would not vary drastically.

7. Errors in both genetic and phenotypic parameters

The reduction of the efficiency of the index with erroneous heritabilities, genetic and phenotypic correlations is illustrated by line 3 in Fig. 4. It basically follows the pattern of the changes in efficiency of the indices with biased phenotypic correlations. If errors are limited within the −30 to +20 percent level in all parameters, the selection indices would excell any single trait selection methods despite faulty coefficients of indices.

The over all results clearly demonstrated the reliability of the Smith-Hazel selection index method. Although errors in phenotypic correlations reduced the efficiencies of the indices rapidly when the magnitude of errors exceeded +/−20 percent in this study, reliable estimates of them are relatively easily obtainable. Pease[10] reported that the expected loss in efficiency of the selection index for swine is usually less than five percent and certainly less than 10 percent when the conceivable combined errors are made. Those facts also support the robustness of a selection index under

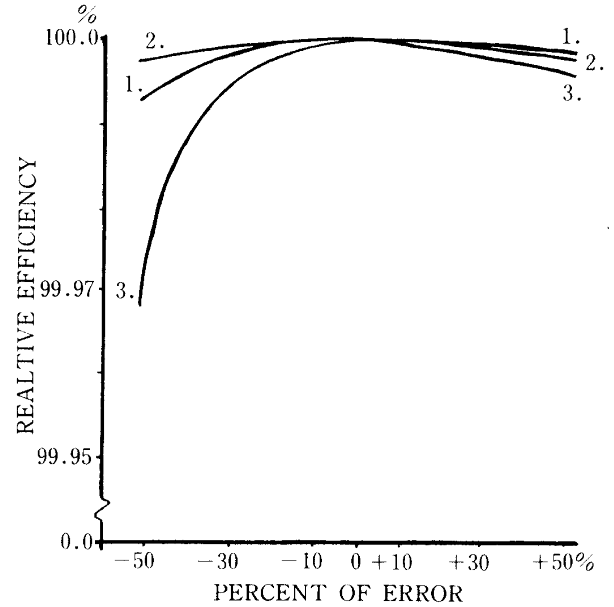

Fig. 5. The relative efficiency when economic weights are affected by errors. The lines 1, 2 and 3 represent the changes in the efficiency when errors are in the economic weights of ear number, of yield and of ear number and yield, respectively.

the violation of the 'error free parameter' assumption as well as the efficiency of a selection index method.

Summary

One of the drawbacks of the Smith-Hazel's selection index is the uncertainty of the effects on the efficiency when one of its assumptions — correct parameters must be known — is violated. Therefore, in this study, the robustness of the index under the violation of the assumption was investigated. First, the reference index for corn (*Zea mays* L.) was constructed. Then, the efficiencies of indices with erroneous parameters (up to +/−50 percent of true values) were compared with efficiency of the reference index. The efficiencies of single trait selection for six yield related characters used to construct the index were also calculated.

The results indicated that, in general, the parameters are over or under estimated even by 50 percent, the selection index still can manage to maintain the 90 percent of the efficiency of the correct index. Although the effects of errors in phenotypic correlations seemed to be larger than those in other parameters, they can be obtained by relatively simple calculations. Therefore, if the population used to estimate the parameters is large enough to minimize sampling errors and the experiment is carefully executed, the chance we obtain grossly faulty estimates would be relatively small. The effects of errors in economic weights were almost negligible. The best single trait selection method was for yield and only 65.8 percent as effective as the correct index.

The overall results of this study agree with those of researchers done by swine breeders[10,20)]

References

1) Comstock, R. E.: Quantitative genetics in maize breeding in Walden, D.B. (eds.), *Maize Breeding and Genetics*. p. 197–206, John Wiley and Sons, New York. (1978)

2) Cunningham, E. P.: The relative efficiencies of selection indexes. *Acta Agri. Scand.*, **19**, 45–48 (1968)
3) Cunningham, E. P.: Multi-stage index selection. *Theor. Appl. Genet.*, **46**, 55–61 (1975)
4) Cunningham, E. P., Moen, R. A. and Gjedrem, T.: Restriction of selection indexes. *Biometrics*, **26**, 67–68 (1970)
5) Elgin, J. H. Jr., Hill, R. R. and Seiders, K. E.: Comparison of four methods of multiple trait selection for five traits in alfalfa. *Crop Sci.*, **10**, 190–193 (1970)
6) Hazel, L. N.: The genetic basis for constructing selection indexes. *Genetics*, **28**, 476–490 (1943)
7) Kempthorne, O. and Nordskog, A. W.: Restricted selection indices. *Biometrics*, **15**, 10–19 (1959)
8) Miller, J. D., James, N. J. and Lytene, P. M.: Selection indices in sugar cane. *Crop Sci.*, **18**, 369–372 (1978)
9) Ordas, A. and Stucker, R. E.: Effect of planting density on correlations among yield and its components in two corn populations. *Crop Sci.*, **17**, 926–929 (1977)
10) Pease, A. H. R.: The pig industry development authority's selection indices. *Anim. Prod.*, **9**, 279–280 (1967)
11) Pesek, J. and Baker, R. J.: Desired improvement in relation to selection indices. *Can. J. Plant Sci.*, **49**, 803–804 (1969)
12) Robinson, H. F., Comstock, R. E. and Harvey, P. H.: Estimates of heritability and the degree of dominance in corn. *Agron J.*, **41**, 353–359 (1949)
13) Robinson, H. F., Comstock, R. E. and Harvey, P. H.: Genotypic and phenotypic correlations in corn and their implications in selection. *Agron. J.*, **43**, 282–287 (1951)
14) Robinson, H. F., Comstock, R. E. and Harvey, P. H.: Genetic variances in open pollinated varieties of corn. *Genetics*, **40**, 45–60 (1955)
15) Sato, M. and Nordskog, A. W.: On estimating components of genetic variance in diallel matings. *Bri. Poult. Sci.*, **18**, 669–704 (1977)
16) Silva, J. C.: Genetic and environmental variance and covariances estimated in the maize (*Zea mays* L.) variety, Iowa stiff stalk synthetic. Ph. D. Dissertation, Iowa State University, Ames, Iowa (1974)
17) Smith, H. F.: A discriminant function for plant selection. *Ann. of Eugenics*, **7**, 240–250 (1936)
18) Sprague, G. F. and Tatum, L. A.: General vs. specific combining ability in single cross of corn. *Agron. J.*, **34**, 923–932 (1942)
19) Subandi, Compton, W. A. and Empig, L. T.: Comparison of the efficiencies of selection indices for three trials in two variety crosses of corn. *Crop Sci.* **13**, 184–186 (1973)
20) Vandepitte, W. M. and Hazel, L. N.: The effect of errors in the economic weights on the accuracy of selection indexes. *Ann. Gent. Sel. Anim.* **9**, 87–103 (1977)
21) Williams, J. S.: The evaluations of a selection index. *Biometrics*, **18**, 374–393 (1962)