Agricultural Systems 6 (1980–81) 81–91

A REGRET FUNCTION FOR AGRICULTURAL PRODUCTION SYSTEMS

ILAN AMIR* & URI SHAMIR†

* *Faculty of Agricultural Engineering and † Faculty of Civil Engineering, Technion—Israel Institute of Technology, Haifa, Israel*

&

ROBERT S. BROUGHTON

Department of Agricultural Engineering, McGill University, Macdonald Campus, Ste. Ann de Bellevue, Quebec, Canada

SUMMARY

Decision-making under uncertainty is a difficult task, involving risks. Since losses are to be expected although not necessarily through the decision-maker's fault, a regret concept is adopted. A regret concept and a regret function, based on a linear programming model, are suggested, using analysis of the sensitivity of the optimal plan to changes in random prices and in the amounts of available resources and estimates of the probabilities that each of the random variables will drop below a critical value. The use of this regret function, as an aid in decision-making, to evaluate the plan at the planning stage is explained and clarified by examples and a case study.

INTRODUCTION

Three main phases are involved in planning agricultural systems: (a) determination of the state of the system and of the framework within which it is operated; (b) selection of planning data and criteria from the relatively large number of possibilities by a suitable method and (c) building a model or an information structure to enable the decision-maker to take action.

The determination of the state of an agricultural system is very complicated due both to the inter-relationships between its components and to the fact that many of the main production processes are quantitatively unknown. Agricultural systems include many random variables whose planning values have to be determined from probability distribution functions (PDFs), and often even the PDFs are not

Agricultural Systems 0308-521X/80/0006-0081/$02·50

available to the agricultural planner. Furthermore, due to the uncertainties caused by nature, the plan may prove to be unstable in that it will require modification before the growing season is over (Amir & Shamir, 1972).

Operating a system under such conditions carries with it the risk of losses when the execution deviates from the plan. The planner must be aware of the losses which might occur, even though it may happen that he incurs losses through no fault of his own. Moreover, the planner, generally, neither has the privilege of reaching a solution by a trial-and-error process nor of using sampling techniques to test his choice of actions. Therefore, it is desirable for him to be equipped with an additional measure to evaluate the plan at the planning stage. The regret concept, which penalises the decision-maker only for his responsibility for losses, might well express the agricultural planner's situation. In addition, as will be shown, a regret function can be the desirable measure for evaluating the selected optimal plan.

The problems addressed may be summarised as follows.

(1) How can a well defined regret concept be adopted for agricultural systems?
(2) How, and at what stage of the decision-making process, can a regret function be calculated?
(3) Will the calculation of a regret function improve the management of agricultural systems?

The objective of this paper is to attempt to answer these questions.

THE REGRET CONCEPT

Linear programming (LP) is a well known technique for planning multi-activity agricultural systems. Its application requires the formulation of the objective function and the determination of the constraints reflecting the framework of the system and its inter-relationships. At the planning stage, using the LP model, the planner has to set planning values as follows.

$$\max z = \bar{\mathbf{c}}\mathbf{x} \qquad (1)$$

subject to:

$$\mathbf{A}\mathbf{x} \leq \bar{\mathbf{b}} \quad \text{and} \quad \mathbf{x} \geq 0$$

where: z = net benefit;
$\mathbf{x} = \{x_j\}$ = vector of the activities (decision variables), $j = 1, \ldots, n$;
$\bar{\mathbf{c}} = \{\bar{c}_j\}$ = vector of the expected net benefit coefficients, $j = 1, \ldots, n$;
$\mathbf{A} = \{a_{ij}\} = n \times m$ technological matrix, the coefficients of which are the demand for the ith resource for a unit of the jth activity;
$\bar{\mathbf{b}} = \{\bar{b}_i\}$ = vector of the expected levels of the resources, $i = 1, \ldots, m$.

The parameters b_i and c_j are either random (with known PDFs) or unknown random, but PDFs unknown). The components of **A**, a_{ij}, are assumed to be deterministic.

The optimal solution achieved by LP also provides a post-optimal sensitivity analysis by calculating the following two factors: (1) 'shadow prices', which are the contribution to the objective function of the marginal unit of the resources, and (2) ranges of insensitivity, i.e. for every component of c_j and b_i, it provides a range which, as long as the component's value is in between this range, the optimality of the solution is maintained. These ranges are mathematically valid for a change in a single component so long as the values of the other components maintain their original values. As a result, when changes occur, they affect the value of the objective function and may cause losses even though the optimal basis is maintained and the decision-maker cannot act in a better way. Therefore, the decision-maker should not take the responsibility for these losses, which he can only regret.

On the other hand, when the changes in the planning values violate the ranges of insensitivity, they render the solution non-optimal. If, now, the planner does not change his plan, he takes the responsibility for the resulting losses since he could have improved the plan by searching for another optimal solution.

Consequently, the regret concept is defined as follows: 'A regret provides a measure of the uncertainty and the instability of a system by evaluating the maximal losses expected, due to future changes in the data on which the plan has been based, and as long as the optimality of the plan is maintained'.

THE REGRET FUNCTION

Any of the random parameters of b_i or c_j in eqn. (1) may be denoted by θ. The possible values of θ are $(\theta_1, \theta_2, \ldots, \theta_k, \ldots)$. G is the set of actions (plans) available for selection, g_m, is one of its members, and $z(g_m, \theta_k)$ is the value of the objective function resulting from the selection of plan g_m and the variable θ_k. Since the plan has to be selected in advance—before the value of θ materialises—there is a loss if the plan selected is not the one which is optimal for the actual value of θ_k. Regret can be measured as the difference between the actual net benefit and that which would have been obtained by selection of $g^*(\theta)$, i.e.,

$$r(g_m, \theta_k) = z[g^*(\theta_k), \theta_k] - z(g_m, \theta_k) \tag{2}$$

$r(g_m, \theta_k)$ can obviously be computed after θ has taken on its actual value θ_k. It is desired, however, to estimate $r(g_m, \theta)$ for any selected plan g_m. $r(g_m, \theta)$ Can easily be computed whenever the PDF of θ is known, while difficulties exist for cases when the PDFs are not known. To estimate $r(g_m, \theta)$ for unknown PDFs one can use two general probability inequalities (Savage, 1961)—those due to Tchebysheff and to Cantelli.

The expected value of the random variable θ may be denoted by $\bar{\theta}$, and its standard deviation by σ_θ. Without making any assumptions about θ's PDF, Tchebysheff's inequality states:

$$P[(\bar{\theta} - t\sigma_\theta) \leq \theta \leq (\bar{\theta} + t\sigma_\theta)] \geq 1 - \frac{1}{t^2}; \quad \text{for } t \geq 1 \tag{3}$$

This is a lower limit of the probability of the random variable being within the range of $\pm t$ standard deviations from the expected value. Inequality (3) is symmetric around the expected value. From it, one may develop the following one-sided inequality:

$$P[\theta < (\bar{\theta} - t\sigma_\theta)] \leq \xi/t^2; \quad \text{for } t \geq 1 \text{ and } 0 \leq \xi \leq 1 \tag{4}$$

This gives the upper limit of the probability of θ falling below the value $(\bar{\theta} - t\sigma_\theta)$. ξ In eqn. (4) is a measure of the symmetry of the unknown PDF. (For a symmetric PDF, $\xi = 0{\cdot}5$.)

Cantelli's inequality (Savage, 1961) is a one-sided inequality which also holds true for any PDF.

$$P[\theta < (\bar{\theta} - t\sigma_\theta)] \leq \frac{1}{1 + t^2}; \quad \text{for } t \geq 0 \tag{5}$$

Depending on the values of ξ and t one or other of inequalities (4) and (5) will be tighter, i.e. it will give a lower value on the right hand side. Figure 1 shows the regions in the t–ξ plane in which each of the inequalities gives the lower limits.

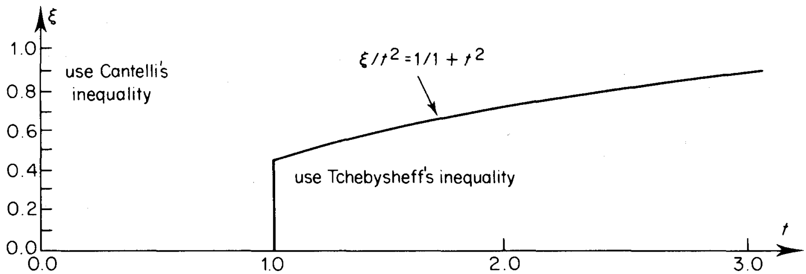

Fig. 1. Regions for using Tchebysheff's and Cantelli's inequalities.

Cantelli's inequality can always be used, regardless of the value of t and whether ξ is known. Thus, when ξ is known, or can be estimated, then Tchebysheff's inequality should be used only for those values of t and ξ for which it gives a better limit.

The above inequalities include the mean and the standard deviation of each of the random variables. In cases where PDFs are not available, it is assumed that these values can, in some way, be estimated. The method of estimation and its accuracy do not cause any theoretical difficulties in the proposed regret function. However, it is obvious that the level of the accuracy affects the stability of the optimal strategy and, thus, may affect the value of the regret function.

Consider the optimal solution of the planning problem, eqn. (1), in which expected values were used. Referring to any random variable θ in the vector **b** or in the vector **c**, let L_θ^* denote the lower limit of the range in which θ must be to maintain optimality. This is the lowest value which θ can take on without causing a change in the optimal plan, $\mathbf{x}^*$, providing all other random variables maintain their original values. These limits are easily obtained from the computer program used to solve the LP problem (for example, the *RANGE* command for sensitivity analysis in MPSX/360 (IBM, 1971)).

The unknown PDF of random variable θ is shown in Fig. 2. Having obtained L_θ^* one can calculate the value of t corresponding to L_θ^*.

$$t = \frac{\bar{\theta} - L_\theta^*}{\sigma_\theta} \tag{6}$$

Using this value, one can compute an upper limit of the probability, shown as the shaded area in Fig. 2, using either Tchebysheff's or Cantelli's inequalities as discussed above.

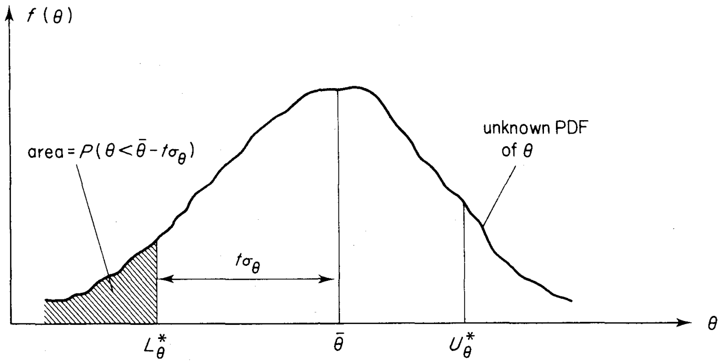

Fig. 2. Graphical representation of a probability measure for an unknown PDF.

U_θ^* is the upper limit of the range in which θ must be, for the solution $\mathbf{x}^*$ to remain optimal. U_θ^* is obtained by the same procedure as L_θ^*. As long as $L_\theta^* \leq \theta \leq U_\theta^*$ and all other variables remain unchanged, $\mathbf{x}^*$ is optimal. (The value of the objective function may change, however, with changes in θ.) When θ is a net benefit coefficient, i.e. $\theta = \bar{c}_j$, then

$$\frac{\partial z^*}{\partial \theta} = \frac{\partial z^*}{\partial \bar{c}_j} = x_j^* \tag{7}$$

and when θ is an input, i.e. $\theta = \bar{b}_i$, then

$$\frac{\partial z^*}{\partial \theta} = \frac{\partial z^*}{\partial \bar{b}_i} = Y_i^* \tag{8}$$

where Y_i^* is the 'shadow price' of the ith constraint in the optimal solution. Thus, for a unit decrease in b_i the value of the objective function decreases by Y_i^*. This parameter is readily obtained in the output of standard LP codes, e.g. in MPSX/360 (IBM, 1971) (this value is denoted as 'Unit Cost' in the output of the *RANGE* command). When maximisation of the objective function is considered, losses will result only if values of b_i and c_j drop below their expected values which were used for planning. When the random variable θ reaches its lower limits of optimality, L_θ^*, the loss in the objective function is:

$$\frac{\partial z^*}{\partial \theta}(\bar{\theta} - L_\theta^*) \tag{9}$$

For $\theta = \bar{c}_j$, expression (9) becomes:

$$x_j(\bar{c}_j - L_{c_j}^*) \tag{10}$$

A similar expression can be defined for $\theta = b_i$, but requires additional explanation. The constraints, which are in the basis of optimal solution, are those which are not entirely used by the optimal activities. That is, a certain amount of their appropriate resources are 'free goods' for which the shadow price, $\partial z/\partial \bar{b}_i = 0$. Therefore, the expression

$$\frac{\partial z}{\partial \bar{b}_i}(\bar{b}_i - L_{b_i}^*) = Y_i^* = (\bar{b}_i - L_{\bar{b}_i}^*) \tag{11}$$

is not zero only for the binding constraints which, mathematically, are not in the optimal basis.

Expressions (10) and (11) give the decrease in the value of the optimal objective function due to changes in the random parameters, under the following restrictions.

(1) Only one parameter undergoes a change, while all others retain their original values.
(2) The change retains the variable within the range ($L_\theta^* \leq \theta$).

In real situations one or both of these restrictions may be violated. Simultaneous changes in several random parameters can be expected (violating the first restriction) to cause a compound loss. This compound loss is, theoretically, smaller than, or equal to, the total loss calculated as a sum of losses due to each parameter separately dropping to its lower boundary, L_θ^*. Furthermore, to compute the compound loss due to specific sets of changes, one would have to solve a new LP for each set and compare the values of the objective function. This method is impractical considering the large number of random variables and the enormous number of possible situations that may occur. Therefore, by assuming that all of the random variables are mutually exclusive, one finds the upper limit of the expected loss. Regarding the first restriction, the regret function calculates the expected loss assuming that each of the random variables is dropped to its lower limit while maintaining the optimality of the solution. Thus, it is consistent with the regret concept previously discussed.

Mathematically, the contribution of each of the random variables to the loss is calculated by eqns. (10) and (11) and is weighed by an estimate of the probability of the particular random variable dropping below its lower limit. The regret function is the summation of all of the contributions as follows:

$$R = E\left(\sum_{j=1}^{n} R_j\right) + E\left(\sum_{i=1}^{m} R_i\right) = \sum_{j=1}^{n} P_j X_j^*(\bar{c}_j - L_j^*) + \sum_{i=1}^{m} P_i Y_i^*(\bar{b}_i - L_i^*) \quad (12)$$

where R is the regret in monetary units (as can be ascertained by using the units of the variables in eqn. (12)), E is the expectation, P_j and P_i are probabilities calculated by eqns. (4) or (5), $X_j^*(\bar{c}_j - L_j^*)$ and $Y_i^*(\bar{b}_i - L_i^*)$ are the maximal loss of every $\bar{c}_j$ and $\bar{b}_i$ within their ranges of insensitivity. When Tchebysheff's inequality is used, then by substituting the relevant expressions, the regret function is:

$$R = \sum_{j=1}^{n} \frac{\xi_j \sigma_j^2}{\bar{c}_j - L_j^*} X_j^* + \sum_{i=1}^{m} \frac{\xi_i \sigma_i^2}{\bar{b}_i - L_i^*} Y_i^* \quad (13)$$

and when Cantelli's inequality is used, then the regret function is:

$$R = \sum_{j=1}^{n} \frac{\sigma_j^2(\bar{c}_j - L_j^*)}{\sigma_j^2 + (\bar{c}_j - L_j^*)^2} X_j^* + \sum_{i=1}^{m} \frac{\sigma_i^2(\bar{b}_i - L_i^*)}{\sigma_i^2 + (\bar{b}_i - L_i^*)^2} Y_i^* \quad (14)$$

DISCUSSION

From eqn. (13) it can easily be seen that regret increases with the variance of the random variable, with the amount of the activity and with the shadow price of the constraint, and decreases with the 'distance' between the expected value of the random variable and the lower limit of optimality. The greater the distance, the more stable the optimal solution, which, of course, diminishes the regret value. These relationships are intuitively appealing and they support the suggested definition of regret.

The other expression of the regret function which is based on Cantelli's inequality, eqn. (14), provides the same behaviour as the one which is based on Tchebysheff's inequality, eqn. (13), i.e. the various components involved cause the resulting regret value to increase or decrease accordingly.

The total expected regret of the planner, expressed by monetary values, should be useful in comparing several possible plans, each of which is the solution of an LP based on some values of the random variable. Although, being intuitively appealing, it is realised that this use of the regret function requires wide and very specific experience which is rather uncommon and takes much time. However, other uses are available and may be obtained by analysing the output of the regret function. By observing the relative magnitudes of the different components of the regret function, one can select those which contribute most to the total regret. For example, if, in the first sum of eqn. (13), one finds a large contribution of some net benefit coefficient, c_j, one may attempt to reduce it. Such a reduction may be achieved by entering into an agreement with the buyer of the product of a certain activity, say a crop. The agreement may reduce the variance considerably (or possibly eliminate it completely) in return for some reduction in the expected net benefit. Similarly, if, in the second sum in eqn. (13), one finds some resource, b_i, which adds a large contribution to the regret, it may be possible to improve the situation. An example might be a case in which b_i is the quantity of rain available in a certain month, and it turns out that its large variance causes it to count heavily in the regret function. The planner might consider using supplementary irrigation. The additional cost of providing the equipment necessary for this irrigation can be weighed against the reduction in regret.

Another example may be the planning of haying (Amir *et al.*, 1978). Haying consists of several processes; cutting, drying, raking, packing, transportation, etc. Selection of the optimal scheme for haying, where the decision variables are the specific haying methods and the appropriate equipment, is accomplished by LP. The number of machines of various types and of working days available significantly affects the amount and quality of hay produced. Variations in the available resources contribute to the regret function in the way explained above. To offset the regret one may enter agreements with sources for the supply of machines and/or manpower. Such agreements require certain expenditures, thereby reducing the final benefit, and this reduction has to be weighed against the reduction in the value of the regret function.

Decisions aimed at reducing regret always cost in some other objectives. Selection of the best policy is a problem of multiple-criteria decision-making. It can be approached by certain methods, which are beyond the scope of this paper (for examples, see Cochrane & Zeleny, 1973).

Thus the relative magnitudes of the components of the regret function help the planner to focus his attention on certain crops, equipment and resources. He can then consider possible steps for reducing the overall regret by actions aimed at reducing individual contributions to the regret.

CASE STUDY

The following is taken from a plan of the production of an irrigated area of land totalling 730 dunams (1 dunam = 0·1 hectare), on a kibbutz farm in northern

Israel. There were 14 different crops from which to choose. The production was constrained by the monthly and yearly quantities of water available from a reservoir, by man power and by crop rotation. The production was formulated as an LP problem to maximise net income. The data entered were estimations of the expectations of both vectors **c** and **b**, where the elements of the matrix **A** were assumed deterministic. For the regret function it was assumed that the standard deviations, σ_j and σ_i, were 15% and 10% of the expectations, respectively. These values were quite arbitrary due to the lack of more accurate estimations. However, in spite of this inaccuracy, the estimations were accepted only for examining the validity of the regret function as an additional tool for the decision-maker.

Resulting from the solution, only five crops were selected for the optimal plan: irrigated and unirrigated cotton, beets, wheat (for rotational considerations) and potatoes.

The optimal objective function was 169400 IL (Israeli Pounds) achieved from the entire land area for the year. The regret function was calculated at the planning stage using Cantelli's inequality (eqn. (14)) since several of the ts were smaller than 1.

Table 1 presents the contributions of the uncertainty to the regret, involved in the *net income coefficients*, c_j; *and Table 2 presents the same for the binding resources* (the contributions of the 'free' resources to the regret are zero, as previously explained).

From these tables one can see the following:

(1) The total contribution to the regret due to the uncertainty of the prices (net income) was approximately 25% only, in spite of the fact that the standard deviations were 15% (potatoes were responsible for 60% of this total).

(2) The total contribution of the resources was 75%, where the main contributors were the labour constraints and the yearly amount of water.

While keeping in mind the inaccuracy of both the expectations and, particularly, of

TABLE 1

COMPONENTS OF THE REGRET FUNCTION ASSOCIATED WITH ACTIVITIES (USING CANTELLI'S INEQUALITY)

Crop	*Planned area,* x_j^* *(dunam)*[b]	*Expected net income,* $\bar{c}_j$ *(IL/dunam)*	*Lower limit,* L_j^* *(IL/dunam)*	*Range* $\bar{c}_j - L_j^*$ *(IL/dunam)*	*Standard deviation,* σ_j *(IL/dunam)*	*No. of* σ_j t_j	*Regret value,* R_j *(IL)*	(%)[a]
Irrigated cotton	360	210·50	208·51	1·99	31·58	0·063	713·55	2·64
Unirrigated cotton	125	115·00	103·35	11·65	17·25	0·675	1000·17	3·70
Beets	55	339·10	198·85	140·25	50·87	2·757	896·84	3·31
Wheat	60	103·35	101·36	1·99	15·50	0·128	117·52	0·43
Potatoes	130	418·48	339·09	79·39	62·77	1·265	3969·50	14·67
$\sum x_j^* =$	730					$\sum R_j =$	6697·58	24·76

[a] The basis is $R = 27054{\cdot}91 = 100\%$ (see Table 2).
[b] 1 dunam = 1000 m^2 = 0·1 hectare.

TABLE 2

COMPONENTS OF THE REGRET FUNCTION ASSOCIATED WITH RESOURCES (USING CANTELLI'S INEQUALITY)

Binding resources	*Units*	*Quantity of the resource,* $\bar{b}_i$	*Lower limit,* L_i^*	*Range* $\bar{b}_i - L_i^*$	*Standard deviation,* σ_i	*No. of* σ_i t_i	*Shadow price,* Y_i^* *(IL/unit)*	*Regret value,* R_i *(IL)*	*(%)*
Water									
June	m^3	23000	19853	3147	2300	1·37	1·13	1236·09	4·57
July	m^3	27000	21284	5716	2700	2·12	0·51	530·57	1·96
August	m^3	27000	23119	3881	2700	1·44	0·72	909·14	3·36
Total	m^3	275000	156875	118125	27500	4·30	0·93	5636·54	20·83
Labour									
July	days	1625	1321	304	162·5	1·87	79·28	5359·50	19·81
August	days	1625	1283	342	162·5	2·10	71·04	5359·50	16·60
October	days	1625	1041	584	162·5	3·59	52·19	4490·88	8·11
							$\sum R_i =$	20357·33	75·24
							$\sum R_j =$	6697·58	24·76
							$\therefore R =$	27054·91	100·00

the standard deviations, the results of the regret function were discussed with the decision-makers of the kibbutz, resulting in the following decisions.

(1) *Regarding the yearly amount of water*. As mentioned above, in this case, the water is stored in a reservoir which is filled by rains during the winter (November to April), while the irrigation season is from May to October. Due to the value of the regret, it was decided to check the reservoir in February and to fill it artificially (by pumping) in case it was not full.

(2) *Regarding labour availability*. This uncertainty was well known to the decision-makers and no concrete steps were taken in advance.

(3) *Regarding potatoes*. It was decided to check very carefully the expectation of the net income and, particularly, the standard deviation by analysing the prices achieved in the past as a function of both the marketing time and of the place. It was also decided to analyse economically the possibility of storing the potatoes rather than selling them fresh.

As one can see, the regret function focuses the attention of the decision-maker on the bottle-necks of his plan, even in the cases where operational decisions are not taken. That is, even without making any decisions due to the numerical results, the planner could be aware of the instability factors of the optimal plan, expressed by the relative regret contributions.

CONCLUSIONS

The regret concept with an adequate mathematical equation is suggested as an additional tool for planning agricultural systems. The lack of sufficient data is met by using probability inequalities and ranges of insensitivity obtained from an LP solution. The regret is computed by a certain state of nature, which is assumed to be the true one at the planning stage, and by the best strategy taken which is the LP optimal solution.

The regret expresses the expectation of the maximal losses as long as the optimality is kept. In other words, if actual events in the future prove that the state of nature is not true, these losses would be the maximum penalty for the need to take a better strategy and to change the plan.

The absolute values of the suggested regret function, *per se*, seem to be of minor importance. However, use of the relative values, in comparison with each other and/or with the total amounts of the regret, has the following advantages.

(1) Focusing the decision-maker's attention on the various instability contributions.
(2) Giving the decision-maker an idea as to how he can reduce these instability factors, even by reducing the expected income.
(3) Since this regret function can be calculated at the planning stage, it enables the decision-maker to act in advance, using the flexibility of most of the multi-activity agricultural systems.

REFERENCES

AMIR, I. & SHAMIR, U. (1972). *Optimal planning and control of an agricultural production system*, Technical report (in Hebrew), Faculty of Agr. Eng., Technion—Israel Institute of Technology, Haifa, Israel.

AMIR, I., ARNOLD, J. B. & BILANSKI, W. K. (1978). An integer programming model for choosing optimal hay systems, Part I: The model, Part II: An application, *Trans. of the ASAE*, **21**(1), 40–54.

COCHRANE, J. L. & ZELENY, M. (1973). *Multiple Criteria Decision Making*, University of South Carolina Press, USA.

IBM (1971). *Mathematical programming system extended (MPSX)*, program number 5734-XM4, White Plains, New York, USA.

SAVAGE, R. (1961). Probability inequalities of Thebysheff type, *J. Research*, *US Nat. Bureau of Standards*, **65**(B3), 212–22.