## 5.2 Application of operations research techniques in crop protection

W.A.H. Rossing

### 5.2.1 *Introduction*

Management of cropping systems and pathosystems requires (as discussed in Section 5.1) appropriate, well-defined decision procedures and adequate information on the state and dynamics of the system. If these conditions are fulfilled, decisions may be made by evaluating all possible options. This is a feasible and operational method if only a limited number of options exists, and few decisions have to be made. For example, a seedling disease may or may not be controlled chemically; if control measures are taken, the pesticide may be applied as seed dressing or after sowing. However, in crop protection it is a recognized fact that many decisions interact. For example, many pests and diseases react to the nutritional status of the crop which, in turn, is determined by cultural measures such as seedbed preparation and fertilization. Traditionally, the experience of the decision maker plays a major role in such cases. Many combinations or options are ruled out in advance, based on experience and knowledge of the behaviour of the system. However, the approach is not transferable and, although the results may be satisfactory, it is not known whether a different combination of decisions would have led to better results.

The rapidly growing research on expert systems employs empirical experience. Here, researchers try to quantify and make explicit the insight and knowledge of the experienced decision maker. The drawback to this approach is that it freezes the knowledge of the currently 'good' farmer and does not help in the development of better and well transferable information. Previous Sections have shown an approach that may lead to improvement in decision making. Information on the effect of various decisions on the behaviour of the system can be obtained by simulation or by experiments, or by combining both ways. The latter seems the best route.

Decision making is a process which continues throughout the growing season. Methods and techniques developed in operations research can be applied to tune managerial actions to objectives. These objectives need not be economic; they may be, e.g. environmental. With most techniques, however, the objectives must be quantifiable, although sophisticated methods are being developed which can handle qualitative objectives.

One optimization technique which has found wide application, including the scheduling of farm operations, is linear programming. Non-linear regression methods originate from operations research. Dynamic programming, discrete event simulation and goal programming are examples of techniques which have been used in agricultural decision problems, although mainly at the research level. In this Section, the basic principles of linear programming (LP), dynamic programming (DP) and discrete event simulation for decision making in crop protection are discussed. After studying this Section, the reader should be able to recognize the structure of the optimization methods discussed here. Formulating a specific problem, e.g. an LP-problem, will require more experience as will the choice of one optimization method over an other. Handbooks, e.g. van Beek & Hendriks (1985), Dannenbring & Starr (1981), Hillier & Lieberman (1980) and Wagner (1979) are advised for further study.

Before proceeding to explain the techniques, the general structure of a decision problem should be discussed.

### *5.2.2 General structure of a decision problem*

In general, a decision problem involves one or more objective functions, decision variables, constraints and a transformation function. The objective function(s) describe(s) the aim of optimization and measure(s) how 'good' a certain combination of decision variables is. The ways in which the decision maker can intervene in the system are represented by the decision variables. Several combinations of decisions may not be feasible, owing to technical or policy considerations. The feasible combinations are described by the constraints. The transformation function describes the way the system evolves under various decision alternatives.

A solution is optimal only within the boundaries of the constraints. These constraints reflect an opinion on the socio-economic situation and on the technical possibilities. As points of view differ and are incomplete a generally optimal solution does not exist. This can be illustrated by an example from the crop protection practice.

When deciding to apply a pesticide, a farmer usually only weighs costs of treatment against costs due to harmful organisms without treatment. Crop husbandry measures, effects on non-target organisms, pesticide residues, dangers to the health of the person applying the chemical and long-term effects on productivity, constitute technical and policy constraints. Each of these represents a decision problem in itself, but with respect to applying a pesticide, they are treated as given facts which cannot be influenced. The farmer's decision problem is then reduced to: what is the optimum timing of pesticide applications to give the highest financial returns (objective function)? The decision variable is 'treatment' with options 'treat' and 'do not treat'. This concept is called 'supervised control'.

In the 'integrated control' concept, less constraints appear in the decision problem. Here, the constraints of the supervised control problem appear as variables in the objective function. Thus, the number of decision alternatives increases. An example is the explicit minimization of effects on non-target organisms in integrated control, as opposed to 'do not spray more than X active ingredient, to prevent excessive effects on non-target organisms' in supervised control. In the former case, the concentration of active ingredients occurs in the objective function, in the latter case in the constraints.

Optimization of decisions in agricultural management may occur at different levels:

1. Crop husbandry measures (e.g. optimal timing of fertilizer applications);
2. Cropping system (e.g. timing of fertilizer application in relation to pests and diseases);
3. Farm (e.g. optimization of the choice of crops);
4. Policy (e.g. optimal farming systems).

The higher the level of integration, the more complex the decision problem, as the number of decision alternatives grows. In principle, management science provides techniques to deal with problems at all levels of integration.

### *5.2.3 Linear programming*

Linear programming (LP) is a general-purpose technique for determining the best allocation of scarce resources. LP-problems are characterized by an objective function (a way of measuring how good an allocation is), a set of decision variables (the way in which scarce resources can be allocated) and a set of resource constraints (limitations placed on the decision variables to reflect the resource scarcity). The objective function and the resource constraints must be linear in the decision variables. This means that a change of one unit in the decision variables results in a constant change in the value of the objective function and in the resource constraint.

LP-problems cannot be solved analytically. They are solved by an algorithm which involves a finite number of operations, the so-called simplex method, developed by Dantzig in 1947, which has found wide application in managerial decision problems.

In this Subsection, two examples will be presented to convey some idea of the type of problem that can be handled by linear programming. A graphic approach and an algebraic approach to solving LP-problems are given. Finally, some advanced applications are considered.

*LP-problem formulation, Example* 1 Consider a farmer who wants to maximize financial returns from two crops, potatoes and wheat. One hectare of potatoes gives two-and-a-half times the financial return of one hectare of wheat. The farmer faces three constraints: he only has 6 hectares of arable land, potatoes may not be grown more than once every two years and, for reasons of diversity, the farmer does not want wheat to cover more than two-thirds of his arable land. The farmer wants to know how many hectares he should put under potatoes and wheat, respectively.

*LP-problem formulation, Example* 2 is taken from animal ecology. A bird may collect food for itself and its nestlings from two isolated areas. One is at 2 minutes distance, the other at 3. The energy needed to visit the two areas is 5 and 10 Joules, respectively. The prey at the first site has an energy content of 25 Joules per individual, the prey item at the other site 40 Joules. The amount of time needed to find and catch a prey item is 2 minutes at the first site and 1 minute at the second. Once a prey is caught, the bird takes it to its nest. Per day, the predator spends no more than 80 minutes hunting and 120 minutes travelling. To satisfy its own basic energy needs, and that of the nestlings, it requires 600 Joules. How many individuals of each prey species must be caught to maximize the net amount of energy gained?

To solve these problems a more formal notation is useful. Example 1 may be reformulated in this way:

maximize $(w = 2x_1 + 5x_2)$ financial return objective Equation 121

where $x_1$ = number of hectares under wheat,
$x_2$ = number of hectares under potatoes

subject to

| | | |
|---|---|---|
| $x_1 + x_2 \leqq 6$ | total area constraint | Equation 122 |
| $x_1 \leqq 4$ | area constraint wheat | Equation 123 |
| $x_2 \leqq 3$ | area constraint potatoes | Equation 124 |
| $x_1 \geqq 0,\ x_2 \geqq 0$ | $x_1$ and $x_2$ cannot be negative | Equation 125 |

As described in the introduction, the problem consists of an objective function (Equation 121), a number of resource constraints (Equation 122 to 125) and decision variables ($x_1$ and $x_2$).

---

**Exercise 77**
Give the mathematical representation of Example 2.

---

*Optimal solution to an LP-problem: the iso-profit line approach* In Figure 78, a graphic representation is given of the LP-problem of Example 1. The range of values of $x_1$ and $x_2$ permitted by the inequalities is indicated by the shaded area in Figure 78a. The set of permitted values of $x_1$ and $x_2$ is called the solution set or the feasible region and consists of the polygon OABCD. The points O, A, B, C and D are referred to as corner points of the solution space. Points on the transects connecting these corner points are referred to as extreme points.

The objective function, w, can assume different values as represented by the broken lines in the graph of Figure 78b. Each line consists of combinations of $x_1$ and $x_2$ which result in the same amount of profit. Note that these 'iso-profit lines'

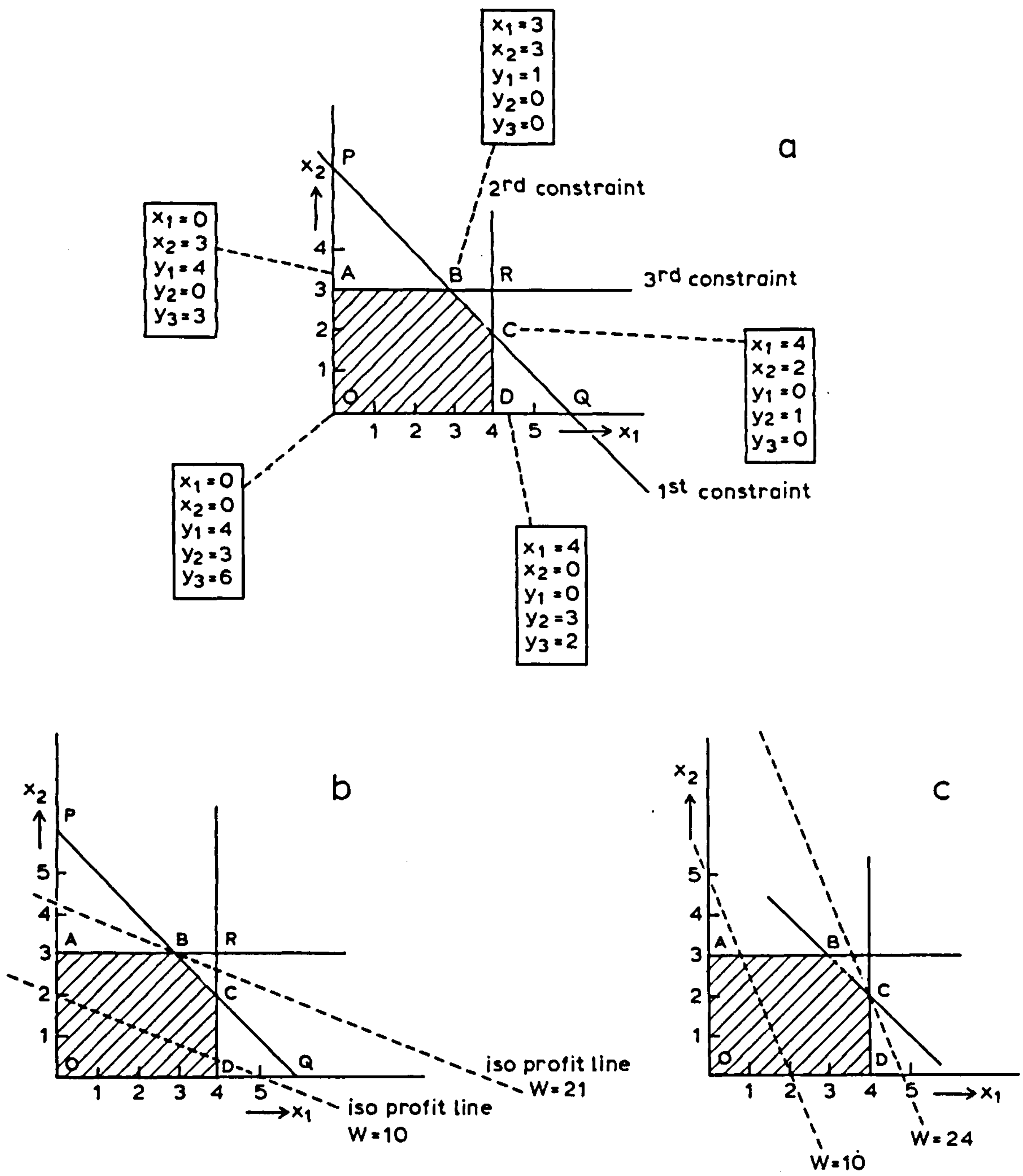


Figure 78. Graphic representation of a linear programming problem. The solution space (shaded) and iso-profit lines (dashed lines).

are perpendicular to the vector $\binom{2}{5}$, which consists of the coefficients of $x_1$ and $x_2$ in the objective function. The combination of $x_1$ and $x_2$ within the feasible region OABCD which is situated on the highest iso-profit line is optimal. In Example 1, the optimal solution is point B, where $x_1 = 3$ and $x_2 = 3$. The objective function value here is 21.

It is not a coincidence that the optimal solution is reached in an extreme point. Only points on the border of the feasible region can be optimal. The explanation can be inferred from Figure 78b. The objective function value can be increased only by moving the iso-profit line 'up' or, more exactly, along the gradient vector $\binom{2}{5}$. Thus, there are two possibilities: the iso-profit line with the highest objective function value either shares one point with the feasible region or coincides with one of its sides. In both cases, the optimal points are extreme points. In the first case the solution is unique, in the second case there are alternative solutions. Thus, when trying to find the optimal solution the search may be limited to the border of the solution space.

It still remains to be proved that point B is optimal. When moving from one point on the border to another the objective function value either increases, decreases or remains unchanged depending on the constraint describing the intermediate intersect. Proceeding from O (with $x_1 = 0$ and $x_2 = 0$) to A in Figure 78, the objective function increases by the marginal contribution of $x_2$ (the contribution of one unit $x_2$) which equals 5, the coefficient of $x_2$ in the objective function. At A ($x_2 = 3$), therefore, the objective function value is 15.

Going from A to B along the constraint line, $x_2$ remains unchanged while $x_1$ increases. The equation describing the line ($x_2 = 3$) shows that $x_1$ can be increased independent of $x_2$ until point B is reached. Here $x_1 = 3$ and $x_2 = 3$. The change in $x_1$ results in a change in the value of the objective function of $2(3) = 6$, so the objective function value is 21. Thus B is preferred to A.

Going from B to C, $x_2$ decreases while $x_1$ increases. As dictated by the equation of the intermediate line ($x_1 + x_2 = 6$), each unit of decrease of $x_2$ is equalled by an increase in $x_1$ of one unit. Thus, each unit decrease of $x_2$ results in an increase of the objective function value of $2(1) + 5(-1) = -3$. When moving from B to C, the objective function value decreases. As the objective function is always linear in LP-problems, it can be concluded that any other point is inferior to B, since A and C are inferior.

---

**Exercise 78**

Repeat this line of reasoning with the objective function $w = x_1 + x_2$.

---

Apart from the unique and alternative solutions illustrated above, two other classes of solutions exist, as illustrated in Figures 79 and 80. In Figure 79, the problem has an unbounded solution set. The objective function value can be made arbitrarily large, while still satisfying the constraints. In a practical setting, this situation may exist for a range of values of a variable until another constraint is reached. In Figure 80, a problem is depicted for which no feasible and, therefore, no optimal, solution exists. As an example, consider the case where a certain minimum amount of pesticide is needed to control a pest while this amount exceeds the maximum level tolerated by beneficial insects.

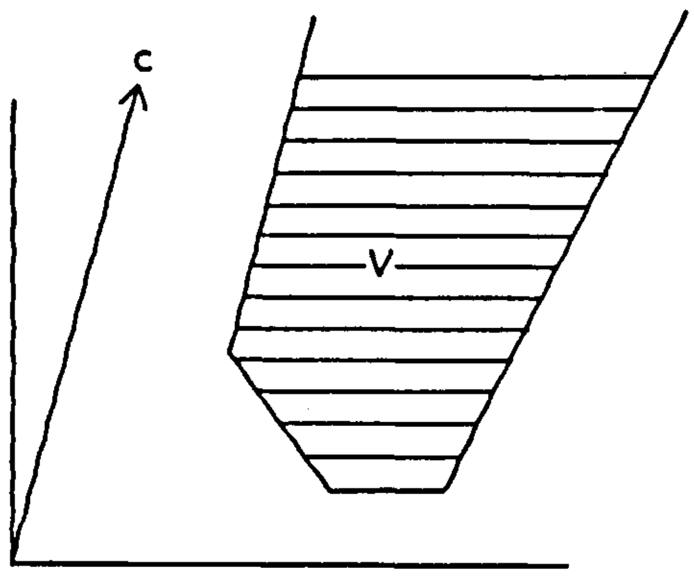


Figure 79. Graphic representation of a linear programming problem with an infinite solution. V is the solution space, c the vector of coefficients of the objective function. (Source: van Beek & Hendriks, 1985).

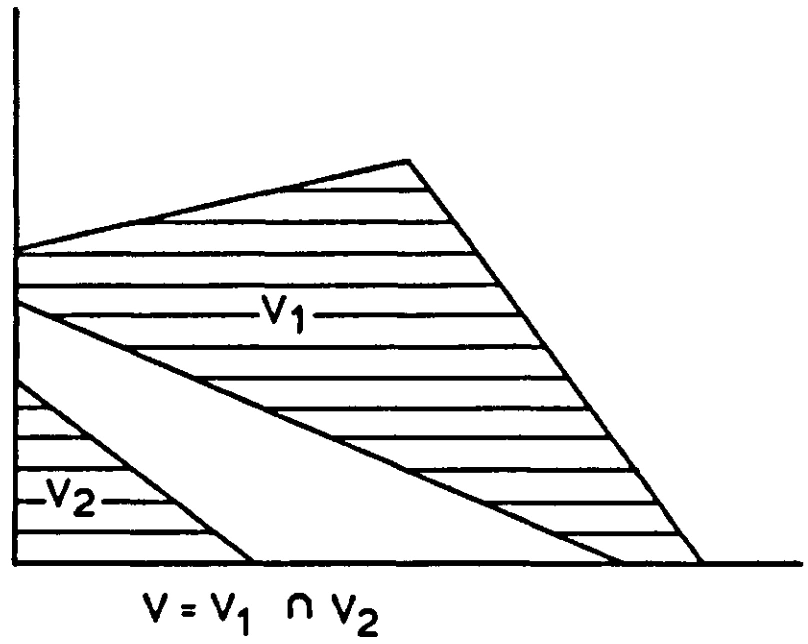


Figure 80. Graphic representation of a linear programming problem without a feasible solution. (Source: van Beek & Hendriks, 1985).

*General algebraic form of an LP-problem* Two forms of LP-problems are distinguished: the standard form and the canonical form. An LP-problem in the standard form can be written as:

$$\text{maximize } (w = c_1x_1 + c_2x_2 + \ldots + c_nx_n) \qquad \text{Equation 126}$$

subject to

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m ,$$

$$x_1 \geqq 0, x_2 \geqq 0, \ldots, x_n \geqq 0$$

Here $c_j$, $a_{ij}$ and $b_i$ ($i = 1,2,..,m; j = 1,2,\ldots,n$) are fixed and the decision variables $x_i$ are to be determined.

In the canonical form the LP-problem is

$$\text{maximize}(w = c_1x_1 + c_2x_2 + \ldots + c_nx_n) \qquad \text{Equation 127}$$

subject to

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &\leqq b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &\leqq b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &\leqq b_m
\end{aligned}$$

$$x_1 \geqq 0, x_2 \geqq 0, \ldots, x_n \geqq 0$$

A number of transformations exists by means of which any LP-problem may be reformulated in its equivalent standard or canonical form. For example, 'maximize (w)' is equivalent to 'minimize $(-w)$', and multiplying the objective function by a scalar $k \neq 0$ does not change the optimal solution. Of course, if $k < 0$, a maximization problem becomes a minimization problem and vice versa. An important transformation is: $\Sigma a_{ij}x_j \leqq b_i$ is equivalent to $\Sigma a_{ij}x_j + y_i = b_i$ (and vice versa), where $y_i \geqq 0$. The variable $y_i$ is called a slack variable as it removes the slack in the constraint. It appears in the objective function with coefficient 0.

---

**Exercise 79**
Apply the transformation rules to derive the standard form of Example 1 and Example 2.

---

*Optimal solution of an LP-problem: algebraic analysis and simplex algorithm* The iso-profit line method with evaluation of extreme border points is only applicable in the case of a two-dimensional LP-problem. For problems of higher dimensionality, the simplex algorithm has been developed. In spite of its name, the method is too complex to be dealt with in detail here. In order to appreciate the hurdles involved in the technique, an algebraic solution to the example in Equations 121–125 will be examined.

Applying the transformations of the previous section, the LP-problem can be rewritten in the standard form:

$$\text{maximize}(w = 2x_1 + 5x_2 + 0y_1 + 0y_2 + 0y_3) \qquad \text{Equation 128}$$

subject to

$$x_1 + y_1 = 4 \qquad \text{Equation 129}$$

$$x_2 + y_2 = 3 \qquad \text{Equation 130}$$

$$x_1 + x_2 + y_3 = 6 \qquad \text{Equation 131}$$

$$x_1 \geqq 0, x_2 \geqq 0, y_1 \geqq 0, y_2 \geqq 0, y_3 \geqq 0 \qquad \text{Equation 132}$$

The simplex algorithm starts with a feasible solution, checks whether it is optimal, identifies a better solution if optimality is not yet reached and stops if it is. A reasonable initial feasible solution seems to be $x_1 = 0$ and $x_2 = 0$ and the slack variables at their maximum (why?) values $y_1 = 4$, $y_2 = 3$, $y_3 = 6$. As the coefficients of the slack variables in the object function are zero, $w = 0$.

The positive coefficients of $x_1$ and $x_2$ in the objective function imply that increasing either $x_1$ or $x_2$ results in an increase of w. The marginal contribution of $x_2$ is greater. Thus, for the purpose of maximizing w, increasing the value of $x_2$ is most attractive.

To which value can $x_2$ be increased while still satisfying the constraints? From Equation 130 it can be seen that $x_2$ may be maximally 3, otherwise $y_2$ becomes negative. According to Equation 131, $x_2$ may be maximally 6. Thus, the maximum value of $x_2$ is 3 according to the most limiting constraint. If $x_2 = 3$, Equation 131 dictates that $y_3 = 3$. The value of $x_1$ is left at zero, $x_1 = 0$, therefore $y_1 = 4$. The second solution, still feasible, is therefore $x_1 = 0$, $x_2 = 3$, $y_1 = 4$, $y_2 = 0$, $y_3 = 3$ and the objective function value is $w = 15$.

How can this be interpreted graphically? The initial solution involved point O in Figure 78a. The value of the slack variables designated the shortest distance along the $x_1$- and $x_2$-axis from O to each of the constraints: the distance to constraint Equation 129 was 4 ($y_1 = 4$), etc. Next, $x_2$ was increased. As can be seen in Figure 78a, values of $x_2$ greater than 3 are no longer within the solution space. Therefore, the next feasible solution is point A. With $x_2 = 3$ the slack in the first constraint (Equation 129) has not changed and is represented by the intersect AR, whereas the slack in the second constraint (Equation 130) has been eliminated. The slack in the third constraint (Equation 131) is reduced to 3, represented by the intersects AP and AB.

As a next step in the optimization, a criterion is needed to judge whether increasing the value of any other variable results in an increase in w. As $x_2 = 3$, this variable can be eliminated from the constraints by making it explicit in Equation 130 and using it as a so-called pivot to remove $x_2$ from other equations. This results in:

$$x_1 + y_1 = 4 \qquad \text{Equation 129}$$
$$x_2 = 3 - y_2 \qquad \text{Equation 133}$$
$$x_1 + y_3 = 3 + y_2 \qquad \text{Equation 134}$$

Substituting $x_2$ described by Equation 133 into the objective function (Equation 128) yields:

$$\text{maximize}(w = 2x_1 + 15 - 5y_2 + 0y_1 + 0y_2 + 0y_3) \qquad \text{Equation 135}$$

The positive sign of the coefficient of $x_1$ in Equation 135 indicates that increasing $x_1$ from its original value ($x_1 = 0$) can result in a solution superior to the previous one. Therefore, the solution found so far is not optimal.

To what value can $x_1$ be increased while still satisfying the constraints? From Equations 129 and 134 it can be inferred that $x_1$ may not be made larger than 3, otherwise $y_2$ and $y_3$ become negative. By making $x_1$ explicit in Equation 134 and using it as a pivot, $x_1$ is eliminated from the other equations:

$$y_1 - y_3 = 1 - y_2 \qquad \text{Equation 136}$$
$$x_2 = 3 - y_2 \qquad \text{Equation 133}$$
$$x_1 = 3 + y_2 - y_3 \qquad \text{Equation 134}$$

From these equations it can be seen that if $x_1 = 3$ and $x_2 = 3$, it follows that $y_1 = 1$, $y_2 = 0$ and $y_3 = 0$.

Substituting $x_1$ (Equation 134) and $x_2$ (Equation 133) into the objective function (Equation 128) yields:

$$\begin{aligned} w &= 2x_1 + 5x_2 + 0y_1 + 0y_2 + 0y_3 \\ &= 2(3 + y_2 - y_3) + 5(3 - y_2) + 0y_1 + 0y_2 + 0y_3 \\ &= 21 - 3y_2 - 2y_3 + 0y_1 + 0y_2 + 0y_3 \end{aligned} \qquad \text{Equation 137}$$

The coefficients of all variables are zero or negative. Thus, no more improvement can be expected by increasing the value of any variable, and the optimal solution found is: $x_1 = 3$, $x_2 = 3$, $y_1 = 1$, $y_2 = 0$, $y_3 = 0$.

As shown in Figure 78b, the optimal solution coincides with point B. The value of $y_1 = 1$ indicates that not all the 'room' available according to constraint 1 (Equation 129) is used, as represented by the intersect BR. The other slack variables have assumed the value 0, indicating that constraints 2 and 3 (Equations 130 and 131) are exactly satisfied.

The steps made in the simplex algorithm can be summarized as follows:

1. The LP-problem is transformed into its standard form. In the example, this was done by adding slack variables.
2. A feasible solution is determined. In this case we chose the origin. Other initial solutions may be appropriate.
3. The optimality of the solution is checked by examining the potential changes in the objective function value resulting from changes in each of the variables.
4. A promising adjacent corner point is selected. Note that this involves changing variables with a value of zero into variables with a value greater than zero, and vice versa. In the example, $x_2$ was changed from 0 to 3 in the first step, while $y_2$ decreased from 3 to 0. Variables with a value of zero are called non-basic, the others are called basic variables. The fundamental theorem of linear programming states that the number of basic variables will always equal the number of constraints. The most promising variable, i.e. the one with the highest coefficient in the objective function, is made basic. The variable to become non-basic is the basic variable in the most limiting constraint (check this in the example).
5. The optimality of the solution is checked. If the solution is optimal, stop. If not, return to step 4.

The different solutions can be clearly arranged in a so-called simplex tableau.

The optimality criterion is also represented in the tableau. For advanced treatment of this subject the reader is referred to handbooks.

---

**Exercise 80**
Solve the foraging problem described in Example 2.

---

*Post-optimal analysis* A farmer may be interested not only in the optimal solution but also in the conditions under which it holds. These are investigated in the post-optimal analysis. Reconsider the problem of the previous section. The optimal solution was $x_1=3$ and $x_2=3$. From Figure 78b the consequences of relaxing the constraints for the optimal solution can be seen. By making constraint 1 (Equation 122) less limiting, the optimum moves to R along the third constraint line. Relaxing constraint 3 (Equation 124) moves the optimum to point P along the first constraint line. Relaxing constraint 2 (Equation 123) does not affect the optimal solution, as, with optimality at point B, there is still slack with respect to this constraint.

If the ratio of coefficients of $x_1$ and $x_2$ in the objective function were 5:2 (instead of 2:5), the iso-profit lines would be perpendicular to the vector $\binom{5}{2}$ and the optimum would be point C, as can be seen in Figure 78c. If the coefficients were equal, the highest iso-profit line would coincide with the transect BC (check this). In this case, there is an infinite number of alternative optimal solutions.

---

**Exercise 81**
Identify the range of ratios of the coefficients of $x_1$ and $x_2$ for which the solution $x_1 = 3, x_2 = 3$ is optimal.

---

*Goal programming* In most applications, decision problems consist of more than one objective. If absolute weights can be attached to each objective, the multiple objective problem may be transformed into a single objective problem. For example, if the problem is to maximize the hectarage of potatoes and wheat on a farm, subject to a number of constraints, the preference of crops may be expressed in their price per unit area. Thus, the problem is transformed into maximization of financial output of potatoes and wheat.

However, for many objectives, only a priority order may be distinguished rather than a quantifiable priority of one over the other. An example is profit maximization on a farm versus conservation of scenic elements. Goal programming is a method of dealing with this type of problem.

The method requires that the objectives are placed in priority order. Starting with the objective of highest priority, the method attempts to satisfy each goal, or,

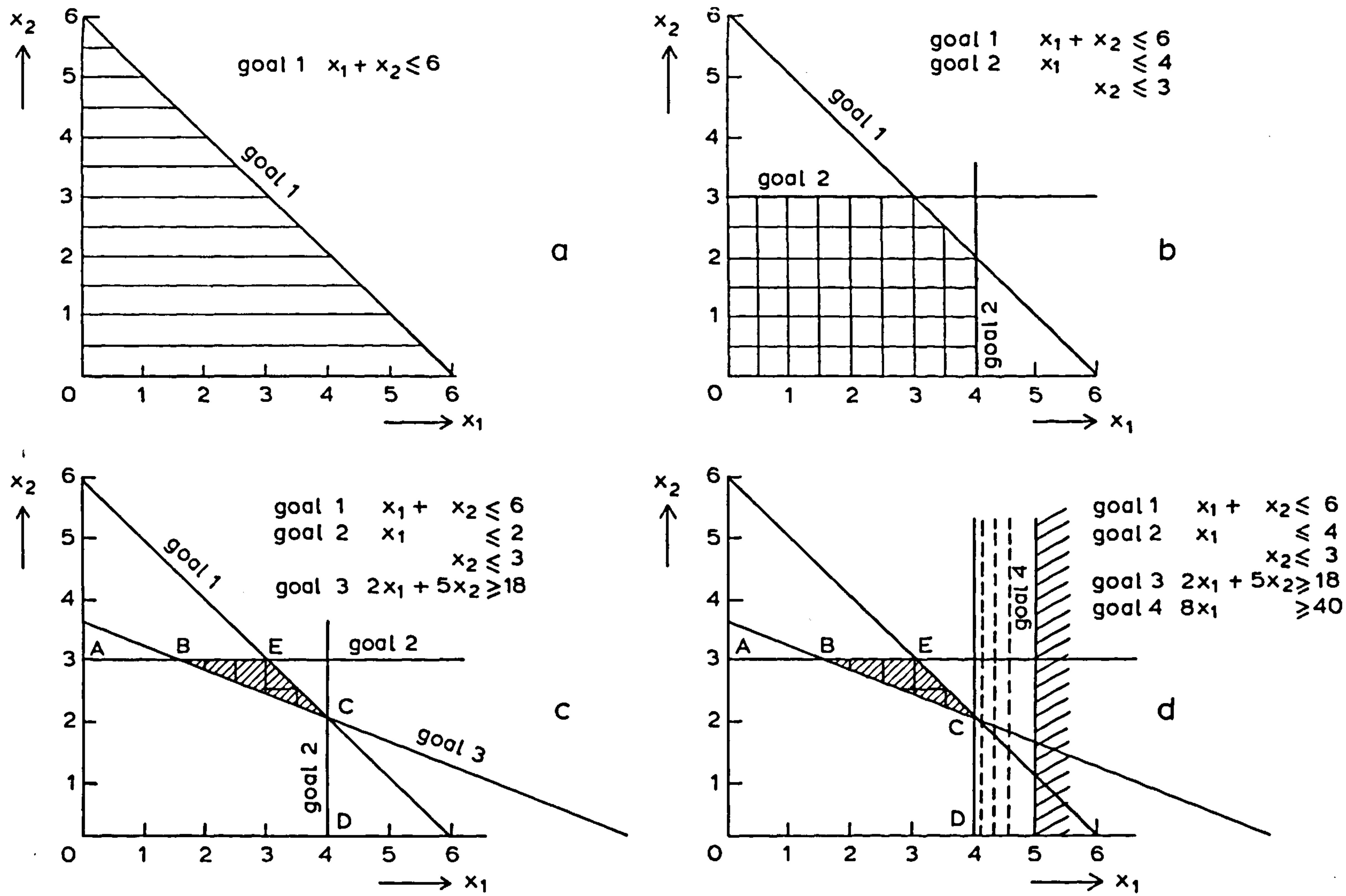


Figure 81. Graphic representation of a goal programming problem. For explanation see text.

failing that, to minimize the undesirable deviations. In this way, a solution can be found that will minimize the amount of underachievement for any goal that cannot be met, without worsening the achievement of any higher-priority goal. Here too, the simplex method can be applied. The following example illustrates the method.

Assume the farmer of Example 1 also wants to produce at least 40 t of wheat to meet a business agreement. His yield expectation for wheat is 8 t $ha^{-1}$. His goals in order of priority are (1) use maximally 6 hectares of land, (2) obey the constraints with respect to rotation, (3) profit at least equal to 18 and (4) produce at least 40 t of wheat.

This two-dimensional goal programming problem can be solved graphically by repeatedly solving the optimization problem, each time adding an extra goal. Thus, the solution space remains unchanged or, if new goals turn out to be constraints, is decreased. Eventually, a goal may be added that cannot be met by any of the solutions satisfying the higher-priority goals. The solution, feasible with respect to the higher-priority objectives, which deviates least from the unsatisfied goal, is designated as being the optimal solution to the problem.

In Figure 81, this approach is illustrated. After having drawn in the first three goals (Figure 81a, b, c), the solution set with corner points BCE is found. The fourth goal (production $\geqq$ 40) cannot be achieved at any of the points in the solution space. Now a solution has to be found that minimizes the deviation from the fourth constraint and is still feasible. The dashed lines in Figure 81d represent combinations of the decision variables $x_1$ and $x_2$ that deviate to the same amount from the fourth constraint. Point C is the first solution encountered which is feasible with respect to the first three goals. Thus, the optimal solution, point C, is to grow 2 ha of potatoes and 4 ha of wheat. The profit is 18 and the deviation from the 40 t level is $1 \cdot 8 = 8$ t.

### *5.2.4 Dynamic programming*

Dynamic programming (DP) is a technique which efficiently determines the optimal policy in problems with separate but related decisions in a set of sequential time periods. DP is generally compatible with pest management models where decisions are made sequentially. The models may be dynamic, non-linear or stochastic. Because of the ability to handle these types of models, DP is a more suitable tool for problems involving timing of chemical applications than LP.

First, the principle of DP will be explained using a deterministic example. It will become clear why DP is an efficient method. Next, a stochastic problem will be treated.

*Shortest route in a network* DP-problems can often be formulated as shortest route problems: finding the shortest route from one state in a network to another. Figure 82 represents such a network with four levels or decision stages (N = 4).

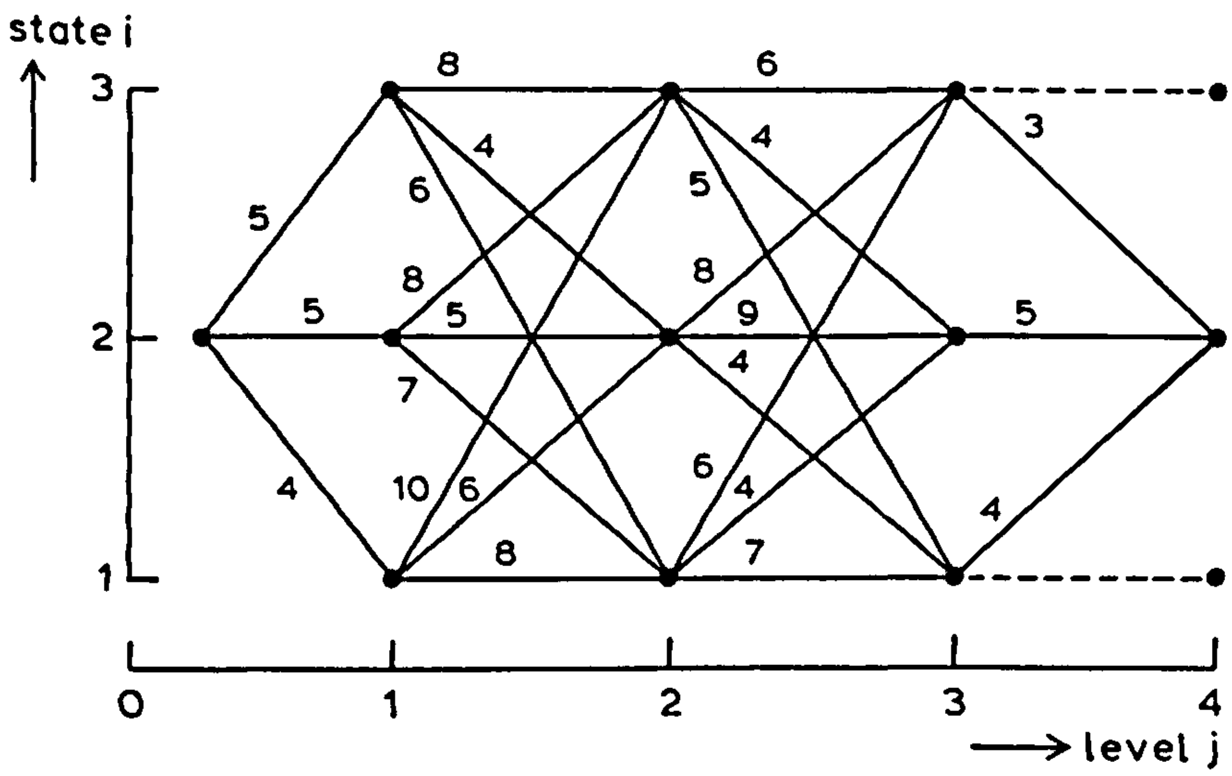


Figure 82. Network representation of a deterministic dynamic programming problem with 4 levels or decision stages. The distance between two nodes or states connected by a branch is indicated.

Each level contains one or more states, nodes in the graphic representation, in which one (the states at the levels 0 and 3) or three decisions (the states on all other levels) are possible. The distances between the nodes vary as indicated in Figure 82. The network may be traversed only from left to right. The problem is to find the shortest route from node 2 at level 0 to the final node 2 at level 4.
Define:

$c_{ik}$ := the distance from node i at an arbitrary level to node k, one level higher
$V_j(i)$ := the shortest route from node (state) i at level j to the final node.

The function V is called the value function. The DP-algorithm states:

$$V_4(2) = 0$$

$$V_j(i) = \underset{k}{\text{minimum}}\,(c_{ik} + V_{j+1}(k))$$

The second line states that, starting in state i at level j, the decision should always be such that the distance to state k at level j + 1 ($c_{ik}$) plus the minimal distance from state k at level j + 1 to the final state ($V_{j+1}(k)$) are minimal. By fixing the value of V at the final (fourth) level and carrying out the calculations starting at the highest level, the algorithm is complete. This procedure can be illustrated using the example of Figure 82.

1. Level 4
   The distance from state 2 at level 4 to any subsequent level and state is zero:
   $V_4(2) = 0$

2. Level 3

Determine the smallest distance from each of the states at level 3 to each of the next states at level 4:

$$V_3(1) = \min_{2}(c_{1,2} + V_4(2)) = 4$$

$$V_3(2) = \min_{2}(c_{2,2} + V_4(2)) = 5$$

$$V_3(3) = \min_{2}(c_{3,2} + V_4(2)) = 3$$

3. Level 2

Determine the smallest distance from each of the states at level 2 to each of the next states at level 3:

$$V_2(1) = \min_{1,2,3}(c_{1,1} + V_3(1), c_{1,2} + V_3(2), c_{1,3} + V_3(3))$$
$$= \min(7 + 4, 4 + 5, 6 + 3) = 9 \quad \text{via state 2 or 3}$$

Here, the values of the value function $V_3(i)$ which were calculated at level 3 are used.

$$V_2(2) = \min_{1,2,3}(c_{2,1} + V_3(1), c_{2,2} + V_3(2), c_{2,2} + V_3(3))$$
$$= \min(4 + 4, 9 + 5, 8 + 3) = 8 \quad \text{via state 1}$$

$$V_2(3) = \min_{1,2,3}(c_{3,1} + V_3(1), c_{3,2} + V_3(2), c_{3,3} + V_3(3))$$
$$= \min(5 + 4, 4 + 5, 6 + 3) = 9 \quad \text{via state 1, 2 or 3}$$

4. Level 1

Determine the smallest distance from each of the states at level 1 to each of the next states at level 2:

$$V_1(1) = \min_{1,2,3}(c_{1,1} + V_2(1), c_{1,2} + V_2(2), c_{1,3} + V_2(3))$$
$$= \min(8 + 9, 6 + 8, 10 + 9) = 14 \quad \text{via state 2}$$

$$V_1(2) = \min_{1,2,3}(c_{2,1} + V_2(1), c_{2,2} + V_2(2), c_{2,3} + V_2(3))$$
$$= \min(7 + 9, 5 + 8, 8 + 9) = 13 \quad \text{via state 2}$$

$$V_1(3) = \min_{1,2,3}(c_{3,1} + V_2(1), c_{3,2} + V_2(2), c_{3,3} + V_2(3))$$
$$= \min(6 + 9, 4 + 8, 8 + 9) = 12 \quad \text{via state 2}$$

5. Level 0

$$V_0(2) = \min_{1,2,3} (c_{2,1} + V_1(1), c_{2,2} + V_1(2), c_{2,3} + V_1(3))$$

$$= \min(4 + 14, 5 + 13, 5 + 12) = 17 \quad \text{via state 3}$$

By storing the best decisions and associated value functions, the set of optimal decisions is recorded. Thus, the shortest route from state 2 at level 0 to state 2 at level 4 is via states 3, 2 and 1 on the subsequent levels and has length 17.

The algorithm was developed by Bellman (1957), who described the principle of optimality which was applied above as follows: an optimal set of decisions has the property that, whatever the initial state and decision are, the remaining decisions must be optimal with respect to the outcome which results from the initial decision.

The value of the value function V in the state at the final level was arbitrarily chosen to be zero. If there are more states at the last level, the final decision may be directed to a preferred state by attaching appropriately high (in the case of minimization) or low (in the case of maximization) values to value functions in unpreferred states. For example, states 1 and 3 at level 4 may be defined to have value function values of infinity:

$$V_4(1) = \infty$$
$$V_4(3) = \infty$$

Thus, they will never be included in the optimal solution.

The efficiency of DP becomes evident when comparing the number of operations carried out with the number of operations needed when checking all possible routes. The DP-method needed 21 additions, 3 for each node at levels 0 to 2, and 14 comparisons of 2 figures, 2 for each node at levels 0 to 2. In an 'exhaustive search', $3 \cdot 3 \cdot 3 \cdot 1 = 27$ routes have to be checked. This involves $3 \cdot 27 = 81$ additions and 26 comparisons of 2 figures. In problems with more states and levels, the discrepancy between the two methods grows in favour of DP.

---

**Exercise 82**

Distinguish the basic components of a decision problem (objective function, decision variables, transformation function and constraints) in the DP-problem formulated in the text.

---

*Stochastic dynamic programming* Stochastic dynamic programming concerns the same type of N-step decision problems as deterministic dynamic programming. However, in stochastic DP, the outcome of a decision is not known with certainty in advance. Two or more outcomes may occur, their likelihood described by a probability distribution. In this case, the optimal policy is the series of decisions which minimizes the expected costs. Calculation is recursive, analogous to the deterministic case. However, the number of computations is generally greater, due to calculation of the expected values.

The algorithm must now be described in more formal terms. Define for $k = 1, 2, \ldots, N$:

- $x_k$ := state of the system at level k. The variable $x_k$ is a vector: the state of the system may be characterized by one or more components.
- $d_k$ := describes the decisions to be made between level $k - 1$ and level k.
- $r_k$ := a vector of stochastic variables, the outcome of which is known at level k. The probability distribution of the variables is assumed to be known and the variables are independent. The stochastic vector is indicated by $r_k$, its outcome by $r_k$.
- $T_k$ := the transformation function: the function which describes the evolution of the system from state $x_{k-1}$, decision $d_k$ and outcome $r_k$ to state $x_k$, yields $x_k = T_k(x_{k-1}, d_k, r_k)$.
- $G_k$ := function describing the costs incurred between levels $k - 1$ and k. These are dependent upon $x_{k-1}$, $d_k$ and $r_k$. This yields $G_k(x_{k-1}, d_k, r_k)$.
- $V_k$ := value function on level k. This is the expected value of costs incurred from state $x_k$ at level k to the final level N, if the series of optimal decisions is implemented. $V_k$ is dependent upon $x_k$: $V_k(x_k)$.

The algorithm now is:

$$V_N(x_N) = 0$$

$$V_{k-1}(x_{k-1}) = \underset{d_k}{\text{minimum}} \{E(G_k(x_{k-1}, d_k, r_k) + V_k(T_k(x_{k-1}, d_k, r_k)))\}$$

where E denotes the expected value.

In the case of stochastic dynamic programming, it is less meaningful to attach different values to $V_N(x_N)$, in order to direct the final outcome towards one preferred state, as the final outcome depends on the chance mechanism.

An example: once a week a decision is made on whether to treat a wheat crop against aphids, based upon the number of aphids and the weather forecast. Two types of weather conditions are distinguished: warm and cool. In the first case, the aphid population grows rapidly and a lot of damage is done. If the weather is cool, the opposite occurs. The decision problem can be formulated in terms of a DP-problem, as illustrated in Figure 83.

The state of the system $x_k$ is described in terms of the number of aphids per wheat tiller. The one-dimensional decision vector consists of the decision to spray or not to spray. Stochasticity is introduced into the system by $r_k$, describing the probabilities (Pr) of the two weather types. The values of variables defining the system at level $k - 1$ are given in Table 38, which also represents the

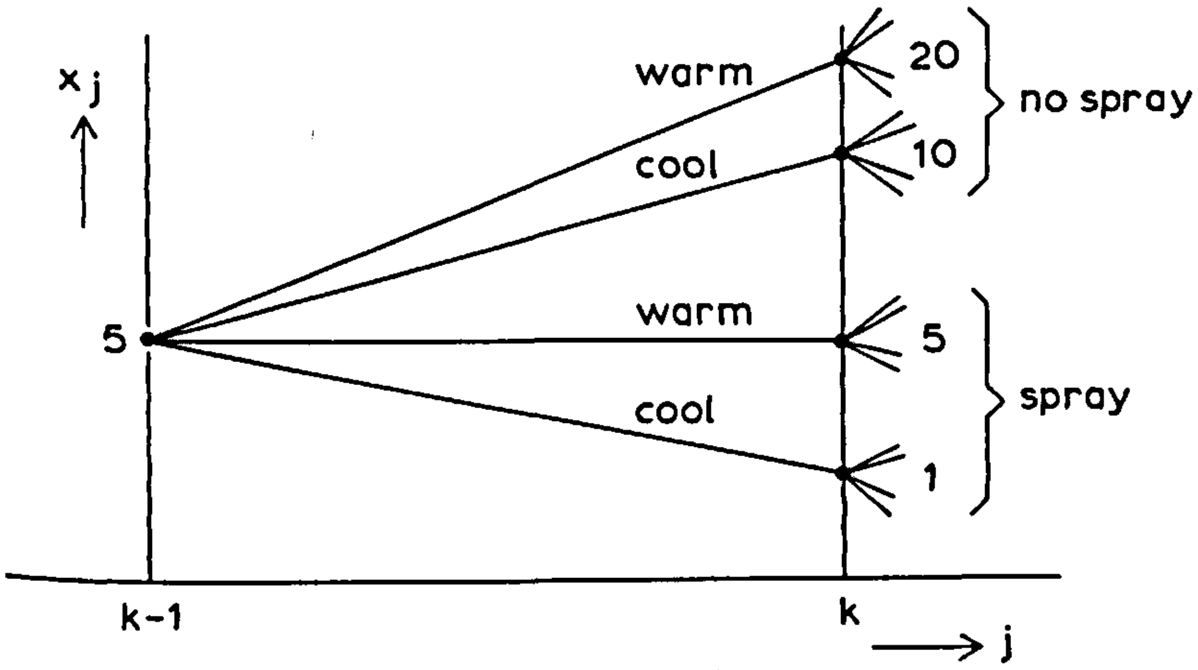


Figure 83. Illustration of a stochastic DP-problem. In each state at each level two decisions may be made: 'spray' or 'no spray.' In each case two type of weather conditions may occur: warm or cool.

Table 38. Chemical control of aphids as an example of stochastic dynamic programming: values of the variable defining the system at level k − 1.

$x_{k-1} = 5$

$$d_k = \begin{cases} 0 \text{ (no spray)} \\ 1 \text{ (spray)} \end{cases}$$

$$r_k = \begin{cases} 1 \text{ (cool), } \Pr(r_k = 1) = 0.8 \\ 2 \text{ (warm), } \Pr(r_k = 2) = 0.2 \end{cases}$$

$T_k$ for $x_{k-1} = 5$:

| $r_k$ \ $d_k$ | 0 | 1 |
|---|---|---|
| 1 | 10 | 1 |
| 2 | 20 | 5 |

$G_k$ for $x_{k-1} = 5$:

| $r_k$ \ $d_k$ | 0 | 1 |
|---|---|---|
| 1 | 35 | 205 |
| 2 | 100 | 225 |

evolution of the system from level $k-1$ to level k. For reasons of clarity, calculation of the value function in only one decision period is discussed. Calculations for the other periods proceed analogously.

The number of aphids in the state vector $x_{k-1}$ is a continuous variable. However, the value function can be calculated only for a finite number of states. Therefore, the value function is calculated for a number of discrete values of $x_k$ and the intermediate values are found by linear interpolation. Suppose $V_k(x_k)$ is known for $x_k = 1$, $x_k = 5$, $x_k = 20$. Then $V_k(10) = V_k(5) + 1/3(V_k(20) - V_k(5))$.

The value function for $x_{k-1} = 5$ can now be calculated:

$$V_{k-1}(5) = \min_{d_k=0,1} \begin{bmatrix} 0.8\overbrace{\{G_k(5,0,1) + V_k(10)\}}^{\text{COOL}} + 0.2\overbrace{\{G_k(5,0,2)+V_k(20)\}}^{\text{WARM}}, & \text{no spray} \\ 0.8\{G_k(5,1,1) + V_k(1)\} + 0.2\{G_k(5,1,2)+V_k(5)\} & \text{spray} \end{bmatrix}$$

With the value function $V_k(x_k)$ known (the calculation proceeds backwards), $V_{k-1}(5)$ is determined.

### 5.2.5 *Simulation analysis of complex decisions*

Both linear and dynamic programming have a number of limitations. Linear programming employs fixed coefficients which may in fact be dynamic. This can be overcome by repeatedly solving the LP-problem with newly calculated coefficients, a time-consuming procedure. Another limitation is the handling of random events. Post-optimal (sensitivity) analysis does not always suffice to estimate the effect of uncertainty in the system. Finally, the model must be linear in the decision variables, a limitation which is hard for many biologists and agronomists to accept, although recently, interesting solutions have been developed (de Wit et al., 1988).

Dynamic programming requires simplified systems with only few state variables, as the computational task soon becomes too large to handle. In the aphid example of Subsection 5.2.4, the number of aphids was the only component describing the state. Other components may include the development stage of the crop, if this interacts with the animal population, or the spraying decisions of previous periods, if residual effects of chemicals are modelled. As a general rule the 'curse of dimensionality' limits the number of state variables to 5 or 6. The inclusion of random events, the number of discrete values of a state component, the complexity of the transformation function and the computer facilities available may alter this figure.

Most of the optimization techniques developed so far are not very suitable for dealing with the combined effects of uncertainty, dynamic interaction between decisions and subsequent events, and complex interdependencies among the variables in the system. When dealing with such a problem, simulation may be a useful tool.

Similar to LP and DP, the system is described in terms of state variables, decision variables and a transformation function. However, simulation does not involve a fixed algorithm. The effect that the choice of decision variables has on model output is evaluated using an objective function. Different series of decision variables can be fed through the model and assessed with respect to efficacy. Thus, simulation is used as a numerical search method, which is very flexible. However, because of the absence of a predescribed solution structure, convergence of subsequent solutions to an optimal solution is not guaranteed as in LP and DP. This is a serious limitation.

System dynamics are simulated with a fixed-time increment or with an event-step increment. In the former case, the procedure is similar to that described in the previous Sections: the state of the system is updated once per time step. In this way, a continuous process is mimicked. In the latter case, the state of the system is updated at time steps, the length of which is dictated by the occurrence of events defined in advance. An example of a discrete event simulation model is described by Tourigny (1985) who simulated the foraging behaviour of *Rhagoletis pomonella*, the apple maggot fly. This insect searches apple trees for host fruits in which to lay eggs. The fruits are patchily distributed, as they occur in clusters. Once a fruit has been accepted for oviposition, it is marked to avoid secondary parasitism. In the model, three events were defined: (1) systematic search of the host tree for fruit clusters, (2) assessment of the fruits within a cluster for oviposition and (3) oviposition.

The occurrence of an event is a random variable. The time to complete an event is also described by a probability distribution. Events can only occur in a logical order: oviposition is not considered before assessment of fruits has occurred.

In the computer model, an event-list defines the state of the system. The event-list gives the last event that occurred, and its completion time. Upon completion of an event, the next event is added to the list, based upon the outcome of the previous event and the random variables generated. The foraging behaviour of individual flies is thus simulated until they emigrate from the host.

Management simulation models may be either deterministic or stochastic. Deterministic models are useful in retrospective analysis of decisions. An example is the assessment of the efficacy of pesticide applications on cereal aphids by British farmers in 1975 and 1977 (Watt et al., 1984). Based on a model of aphid damage, which was correlated to aphid density, the effect of spray timing was evaluated. The evaluation consisted of the repeated execution of a computer program, each time with a different application date of the chemical. Finally, the optimal decisions were compared with the actual decisions and the differences were expressed in financial terms.

Stochastic management models include random phenomena. In crop protection, these may be weather variables, sampling errors, immigration of pests, etc. The term risk analysis is often applied to such decision models. For example, consider a farmer who wishes to minimize expected costs attributable to aphids, subject to the constraint that the risk of incurring high costs should not exceed a certain threshold. If he does not spray, both the risk and the expected costs are high. If he sprays frequently, the risk of aphid outbreaks may be sufficiently low, but the expected costs are again high due to the application costs. Less frequent treatment, and appropriate timing, may satisfy the objective. In this situation, simulation of the decision problem is useful since: (1) the number of decisions is large (a farmer may spray every day of the growing season), (2) the interactions between crop, aphids and environment are complex and (3) as a result of (2) no analytical solution exists for the probability distribution of costs.

The usefulness of a stochastic model in this case depends on the extent to which the stochastic variables influence damage. This example is treated in detail by Rossing (1988).

### 5.2.6 *Final remarks*

Decision-optimization models may be used strategically or tactically. In the first case, the optimal solution for a number of initial conditions is calculated and tabulated or otherwise stored. The decision maker then refers to these tables. In the case of tactical or on-line use, the optimization model is run for one specific set of initial conditions. The approach chosen to disseminate information depends on the problem.

The value of optimization models depends on the quality of the pathosystem description. Poor models of the ecophysiological aspects of population dynamics and damage may yield unrealistic optimal decisions. This is especially true for DP where optimization is carried out over a large planning horizon, i.e. the period of time for which optimal decisions are calculated. Thus, errors when describing the state of the system at the end of the season, affect the first optimal decision calculated. The extent to which this occurs must be evaluated for each case in sensitivity runs with the decision model.

Whether or not further research into certain biological aspects is worthwhile, can be evaluated by sensitivity analysis of the decision model. For efficient planning of biological research activities, it is advisable, therefore, to combine the development of biological models for decision purposes with the development of optimization models.