

ALES: a framework for land evaluation using a microcomputer

David G. Rossiter¹

Abstract. ALES, the Automated Land Evaluation System, is a microcomputer program that allows land evaluators to build their own knowledge-based systems with which they can compute the physical and economic suitability of land map units, in accordance with the FAO's Framework for Land Evaluation. The economic suitability of a land mapping unit for a land utilization type is determined from the predicted annual gross margin per unit area. Increasing limitations result in increased costs of production, decreased yields, or both. Evaluators build decision trees to express inferences from land characteristics to land qualities, from land qualities to predicted yields, and from land qualities to overall physical suitability. A representative model is described.

INTRODUCTION

THE AUTOMATED Land Evaluation System, or ALES, is a computer program that allows land evaluators to build their own knowledge-based systems with which they can compute the physical and economic suitability of land map units in accordance with the FAO's Framework for Land Evaluation (FAO, 1976). This article first discusses the land evaluation problem and the FAO's approach to its solution. It then describes the program. Finally, it illustrates the use of the program with an example.

LAND EVALUATION

Different kinds of land are unequally suited to various uses. Among the differences between land areas are their physical attributes, such as soil characteristics, climate, terrain, and water resources. These are the subject of natural resource inventories such as soil or land system surveys. Lands also differ because of current and past land use, and the social and economic context within which they are used. These are the subject of social geography.

When society and its demands on the land are changing only slowly, the matching of land use to the land itself by trial and error works well. However, in the modern world with its rapid population growth and increasing environmental degradation, changes in land use must be much faster, hence the need for land-use planning at national, regional, local, and farm levels.

One of the main tasks of scientists working in land development is to interpret resource inventories for users and planners of land. These groups must know the suitability of land areas for actual and expected land uses, often in the face of pressing needs. In response to these demands several systems of land classification have been developed over the past 50 years, most notably the land capability classification of the USDA (Klingebiel & Montgomery, 1961).

In view of these the FAO sponsored a series of expert consultations on land evaluation in the early 1970s. These resulted in the landmark *Framework for Land Evaluation*, published in 1976 (FAO, 1976). The framework has since been amplified with the publication of guidelines for land evaluation for several kinds of land use (FAO, 1983, 1984, 1985; Siderius 1984, 1986).

The FAO's method of land evaluation

The FAO's method emphasizes planning new land uses, including crop species, cropping systems, and management methods, in response to changing demands on the agricultural sector especially in developing countries. It has three levels of detail.

At the most general level is the *Framework* (FAO, 1976), which is a method with no content. It tells the evaluator *how* to evaluate. At the middle level are specific *Guidelines*, which provide suggestions for content. These suggest factors that should be considered when evaluating land for a specific purpose and how they might be assessed. The FAO has published guidelines for rainfed agriculture (FAO, 1983), irrigated agriculture (FAO, 1985), and forestry (FAO, 1984). In addition, workshops have been held in which guidelines were discussed for extensive grazing (Siderius, 1984) and sloping areas (Siderius, 1986).

At the most specific level are individual evaluation exercises. Each evaluation applies to a specific project, which may be at the national, regional, or local scale. The evaluator must understand the framework (the most general level) and incorporate suggestions from one or more guidelines (the middle level). However, the details of each evaluation are ultimately decided by the evaluator, taking into account local sources of information, land use objectives, and socio-economic constraints.

The FAO's method contains several key innovations. First is the concept of the *land utilization type* (LUT), a technical term which replaces the common term 'land use'. The LUTs in an evaluation exercise are the land use op-

¹Department of Agronomy, Bradfield & Emerson Halls, Cornell University, Ithaca, NY 14853, USA.

tions, and they include in their definition all aspects of land use that can be affected by differences in land. The land-use system is typically described in detail, including the cropping system, management methods, and socio-economic context. Each land area to be evaluated receives a separate rating for each LUT. Thus there is no intrinsically 'good' or 'bad' land, only land that is more or less suited for each possible use.

Another key idea of the FAO's method is that *economics* must serve as the means of comparing the suitability of diverse LUTs. Relative physical suitability serves only to rank LUTs on an ecological scale. To rank land use options on the scale of desirability for man's use, one must use man's measuring stick, which is money. This statement may seem naive. However, many things that are not normally thought of as having a monetary value (e.g. environmental protection, way of life) may be assigned a social value denominated by money in order to perform an economic analysis.

Perhaps the most important innovation of the FAO's method is the introduction of an intermediate level of describing land, namely *land qualities* (LQ). These are attributes of land that directly influence its suitability for one or more uses. They are more abstract than *land characteristics* (LC), which are the attributes of land that can be measured or estimated. Land qualities may be edaphic (e.g. nutrient availability), climatic (e.g. radiation regime), agronomic (e.g. soil workability), geographic (e.g. location), or related to hazards (e.g. erosion hazard).

Land is described by its land qualities, and the parallel concept of *land use requirement* (LUR) is used to describe land uses. LURs are 'the conditions of land necessary for the successful and sustained practice of [a given LUT]' (FAO, 1984, p. 228). These are matched one-to-one with corresponding land qualities in order to determine the fitness of each land area (which has qualities) for each land use (which has requirements).

Land qualities are generalized attributes of the land that cannot usually be measured directly in the course of routine survey. Instead, each land area is defined by values of a set of land characteristics. The land qualities must then be inferred from some subset of the land characteristic values; these are referred to as the *diagnostic* land characteristics. There are several methods that have been used to express these inferences, including matching tables (FAO, 1983; Sys, 1985), parametric methods (Riquier, 1974), and transfer functions based on process models (Bouma & van Lanen, 1987). *ALES* uses a different method, that of decision trees, which is explained later.

A final innovation of the FAO's method is the concept of *land mapping units* (LMU), which are identifiable areas of the earth's surface, considering all properties of relevance to land use options. The term 'land' is thus used in a very broad sense, and typically includes the concepts of soil, climate, terrain, and current land use. Land units can be map units in a soil or land system survey, consisting of a set of delineations (map polygons) each with the same

identifier. They can also be individual delineations such as fields, or even points on the earth's surface.

Automating the FAO's method

Evaluations based on the FAO's method require many data, involve numerous repetitive calculations or references to tables, and are tedious if many possibilities are to be compared. Manual procedures, both for construction of matching tables or transfer functions and for calculation of suitability, are time-consuming and error prone. Hence an automated procedure is very desirable, and indeed there have been a few attempts to create such, most notably the LECS system in Indonesia (Wood & Dent, 1983; Purnell, 1987). However, this successful system has two serious limitations: (1) it is at the most specific level in the three-level hierarchy of the method, and thus is applicable to a specific scale and area of the world; (2) it is not interactive. To make the FAO's method widely applicable a micro-computer program that could be customized for individual land evaluation projects and their particular circumstances was needed.

THE ALES APPROACH TO LAND EVALUATION

ALES itself is a *framework* within which evaluators build their own models; it does not by itself contain any knowledge. In the terminology of knowledge-based systems (e.g. Waterman, 1985) it is a *shell*, which provides a reasoning mechanism and constrains the evaluator to express inferences using this mechanism. Within this shell evaluators build their own *expert systems*, which can then be used to evaluate land. Thus *ALES* is a computerized realization of the FAO's Framework, and models within the system are computerized realizations of specific evaluation exercises. Figure 1 shows this relationship.

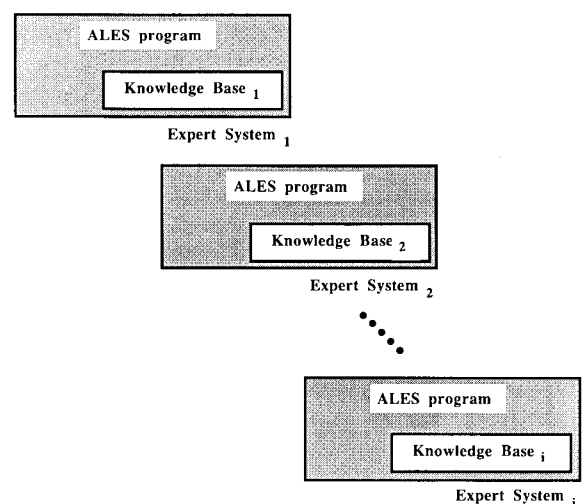


Fig. 1. Relation between *ALES* program and models.

An *ALES* model may be classified, according to the terminology of Burrough (1989), as an *empirical model of reality*, as its empirical decision procedures describe the relation of land attributes to their fitness for specific uses. However, it can also be thought of as a *model of expert judgement*, i.e. the codification in a constrained form of the inferences already present in the mind of the land use expert.

A related issue is how such models should be *validated* (O'Keefe *et al.* 1987). Land evaluation models predict the performance of land areas when used in certain ways. A comprehensive validation would require that all the land units in the area of interest, or at least a representative set of them, be subjected to each land use, with replications both in space (different delineations of each map unit) and time (multiple cropping cycles). Then the actual economic returns, crop yields, and individual land quality ratings could be compared against the model's predictions, using standard statistical techniques. Validation on this scale is patently impossible; indeed, if the resources to conduct such experiments were available, there would be no need for a land evaluation model.

Land evaluation models, then, cannot be validated against reality in the same way as, for example, crop growth models. Instead, they must be considered valid if they accurately reflect the land evaluator's best judgement and explain the way in which the evaluator is making these judgements.

Model building

Land evaluation models are built in the following manner. First, the evaluator builds a preliminary version of the model, by: (1) selecting a few representative land utilization types; (2) expressing these in terms of their most important land use requirements; (3) determining which land characteristics are available to form the basis of evaluation (i.e. making an inventory of available data sources); (4) constructing decision procedures to relate land characteristics to land use requirements; and (5) determining prices and interest rates.

After building the preliminary model the evaluator selects some representative or well understood map units, and he or she collects land characteristic data for these, entering them into the database. Then the evaluator uses the program to compute and display *evaluation matrices*, which show five kinds of ratings for each map unit for each land utilization type, namely: (1) physical suitability subclasses, (2) economic suitability classes, (3) predicted gross margin, (4) expected yields of crops or other outputs, and (5) ratings for single land qualities. The evaluator then compares the ratings displayed in the matrix against values determined by other methods. A particularly useful kind of information is the ranking of map units within each land utilization type. If there are actual yield data then these are compared with the yields predicted.

While viewing the matrix of results the evaluator may

repeatedly press a function key (called the 'why?' key) to follow a *backward chain* showing every step of the reasoning by which a particular map unit was assigned a particular suitability. At each screen, all data, parameters, and decision trees that entered into the current step of the computation may be edited. After any sequence of editing the evaluator may request a fresh recomputation of the matrix. This procedure is repeated until the evaluator is satisfied with the validity of the preliminary model.

Once the preliminary model has been completed, the evaluator may extend it to a wider set of land utilization types. *ALES* allows the evaluator to copy entire land utilization type definitions, and then to modify the copies to reflect differences in, for example, management method, or economic assumptions.

Figure 2 shows the flow in building a model.

Model use

The model can at this point be turned over to clerical staff, who then enter definitions and data for the remaining land units in the evaluation area, using the data entry forms designed by the model builder. They then request of the program a final comprehensive evaluation and printed reports showing the best land areas for each use and the best uses for each land area.

An alternative mode of use of a completed model is the *consultation*. In this mode the program prompts the user to select the correct values of land characteristics for a single land area, and it computes its suitability for one use following a forward chain through all relevant decision trees. This is primarily a teaching tool, but it can be used to evaluate single sites.

THE *ALES* KNOWLEDGE BASE

A *knowledge base* is a structured representation of the facts and inferences needed to arrive at decisions. Figure 3 shows details of the schema used in *ALES* and which must be filled in by the evaluator for each model. Some of the items are single facts, for example, the purchase price of an input, the length of a rotation, or the optimum expected yield. More interesting, however, are the forms in which the evaluator's inferences are represented within the knowledge base.

In the FAO's method, land quality values are determined from the values of a set of diagnostic land characteristics, and the resulting land quality values are in turn combined into physical and economic suitability values. The way in which these combinations are performed must be specified by the model builder. In *ALES* these inferences are typically expressed in the form of decision trees.

Decision trees

Decision trees are hierarchical multiway keys in which the leaves are results (e.g. severity levels of land qualities), and

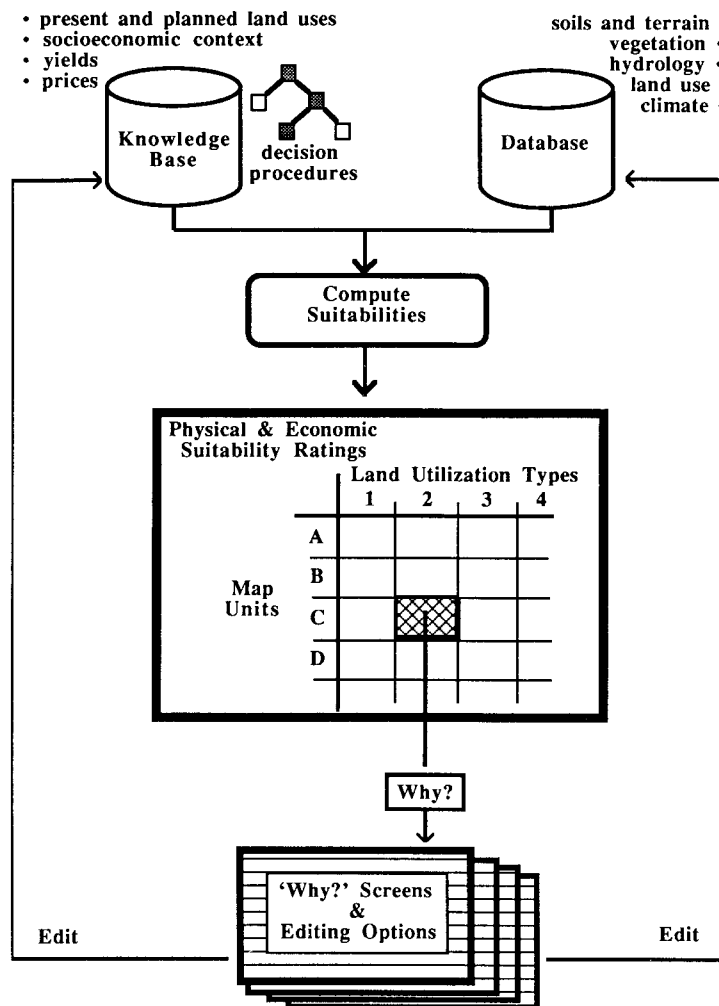


Fig. 2. ALES program flow.

the interior nodes of the tree are decision criteria (e.g. land characteristic values). These trees are constructed by the model builder, and they are traversed by the program to compute an evaluation using actual land data for each map unit. Figure 4 shows a simple decision tree which allows the program to determine the value of the land quality 'likelihood of high fixation of fertilizer phosphorus by iron' from the land characteristics 'ratio of free Fe_2O_3 to clay in the topsoil', 'percentage of clay in the topsoil', 'hue of the topsoil matrix', and 'topsoil structure'. As shown in the example, these decision trees allow missing data to be replaced by other criteria should the model builder accept them.

Decision trees have several advantages for expressing inferences in context. Most notably, both the model builder and user have an explicit representation of the reasoning process used to reach a decision. Trees may be traced by hand or with the aid of the program's 'why?' screens, thus affording insight into the model builder's logic. Trees are typically constructed with the most differentiating criteria

at the highest levels. However, they may also consider factors in some natural order, such as the growth or cultural sequence of a crop. Thus model builders may implicitly communicate some of their thoughts to the user.

Decision trees reason with *classified* data, whether nominal (one of a finite set of unordered choices) or ordinal (a set of naturally ordered choices with an underlying scale of measurement). Data from resource inventories are usually presented as classified values, for example, type of parent material (nominal), soil depth (ordinal, measured in cm) or mean rainfall in the growing season (ordinal, measured in mm). In the second case, the inventory typically reports a range of depths that applies to most of the points in the map unit. In the third case, the amount of rainfall is also presented as a range, which is shown on climate maps as the area between iso-lines.

In some inventories, only single values are given, for example, representative values for soil depth, or a single year's rainfall measured at one station. Therefore, the program allows the model builder to define single-valued vari-

- 1 Evaluations[°]
 - 2 Land Use Requirements[°]
 - descriptive name
 - 2 Inputs[°]
 - units of measurement, purchase cost
 - 2 Outputs[°]
 - units of measurement, selling price
 - 2 Land Utilization Types[°]
 - number of physical suitability classes
 - economic suitability class limits
 - length of rotation
 - physical suitability subclass decision tree[°]
 - 3 Inputs[°]
 - annual & one-time amounts
 - 3 Land Use Requirements[°]
 - number of severity levels
 - severity level decision tree[°]
 - 4 Severity Levels
 - 5 Inputs
 - additional annual & one-time amounts
 - 3 Outputs
 - number of crops per rotation
 - optimum yield
 - proportional yield decision tree[°]
 - 4 ⁺Land Use Requirements
 - 5 Severity Levels
 - limiting yield factor
 - multiplicative yield factor
- 2 Land Characteristics[°]
 - Case: *discrete* land characteristic
 - number of classes, units of measurement, ⁺class limits
 - ⁺land characteristic-to-land characteristic decision tree[°]
 - ⁺commensurate land characteristic from which to infer
 - Case: *continuous* land characteristic
 - units of measurement, upper and lower limits

Underlined entities are their primary definition.
 Numbered entries may be repeated for different entities at the same level.
 Entities marked with [°] may be annotated.

Fig. 3. ALES knowledge base schema.

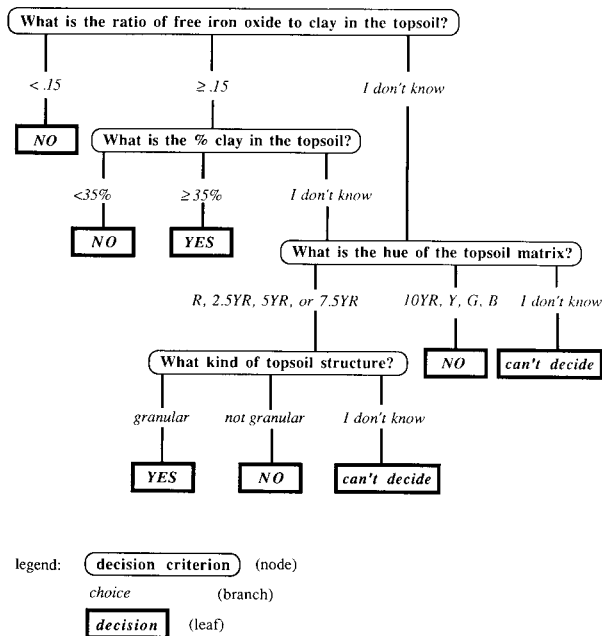


Fig. 4. A decision tree to determine whether there is a likelihood of high fixation of fertilizer phosphorus by iron (after Sanchez, Couto & Buol, 1982).

ables and link them to related classified variables. During computation, single values are assigned to the correct class of the related classified variable.

Decision trees cannot easily express calculations with single-valued variables, for example, predicted soil loss values by a soil-loss equation. If the evaluator wants to compute such values then this computation must be performed outside the program, for example, in a spreadsheet or database program, and the resulting value imported as a single-valued land characteristic.

If many factors must be considered then decision trees become cumbersome, since they grow exponentially with the number of decision criteria. The program therefore allows the use of parametric and limiting yield factors, and the maximum limitation method for physical suitability, as adjuncts to certain decision trees. In addition, the evaluator may introduce intermediate results to break down one tree into several smaller ones.

Annotations

A key feature of a knowledge-based system is its ability to explain its reasoning. Evaluators build decision procedures

and specifications of land utilization types and their components, such as land use requirements, inputs, and outputs. The program allows them to explain their reasoning by entering *annotations* for the components of their models. These are text notes of any length that typically include sources of data, bibliographic citations, and considerations that could not be expressed by means of decision trees. The presence of a note is shown by a symbol on the screen next to the entity name in any context where the entity is referenced (for example in a 'Why?' screen or during a consultation). The user may view the note by pressing a single key.

ECONOMIC ANALYSIS

ALES can be used to build models for physical land evaluation only. However, its full benefit is obtained when both physical and economic evaluations are computed. Models can provide the land use planner with a realistic estimate of the economic suitability of each land unit for each proposed land use. In the FAO's method as published the type of economic analysis, let alone a detailed method for such an analysis, is unspecified. Young & Goldsmith (1977) and Dent & Young (1981) have made important contributions in this regard. *ALES* compares land use options by

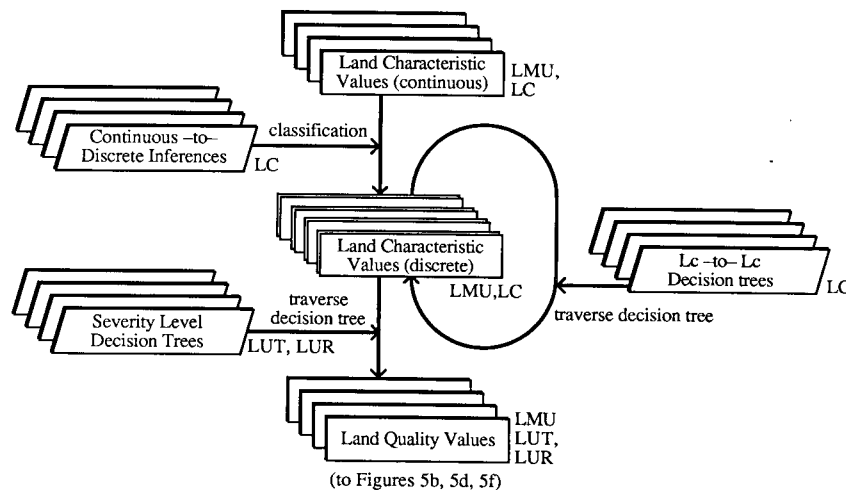


Fig. 5a. Computation of land quality values.

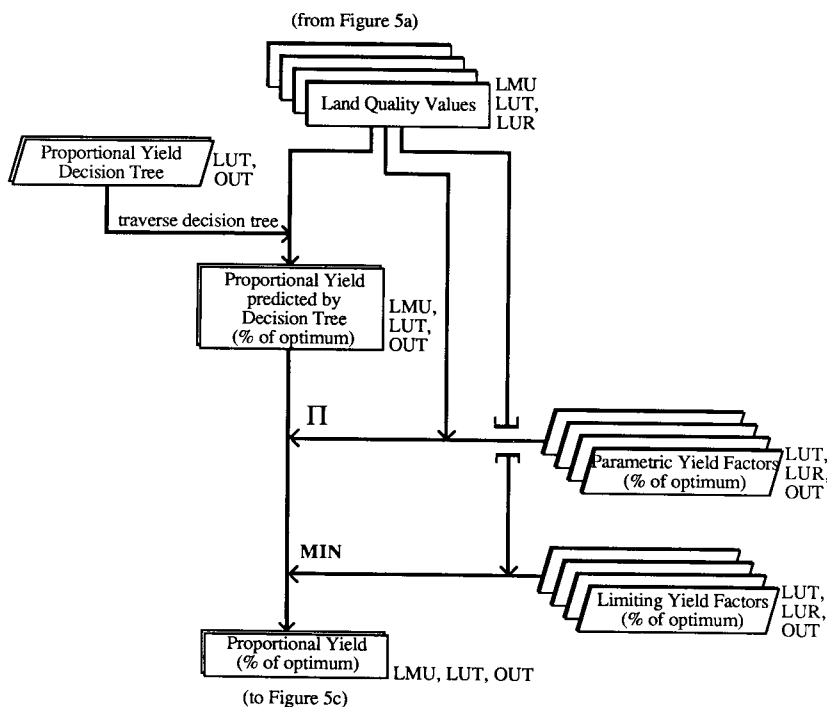


Fig. 5b. Computation of proportional yields.

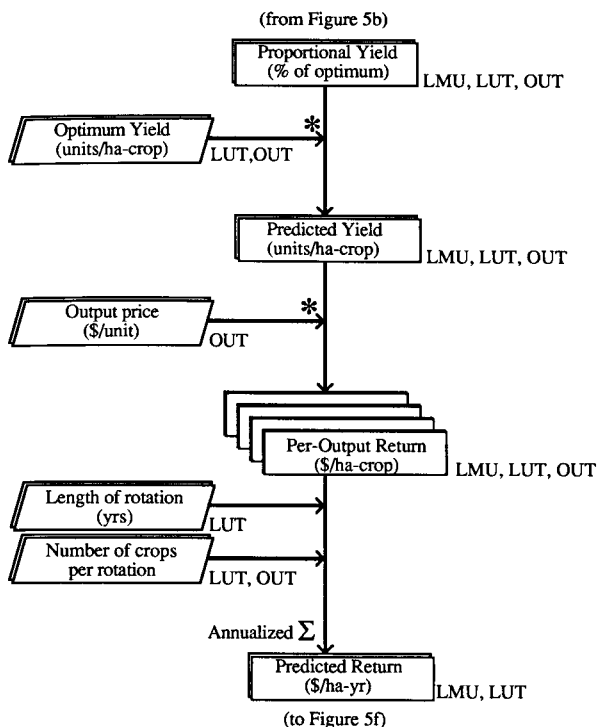


Fig. 5c. Computation of predicted returns.

gross margin analysis. This kind of analysis considers only cash flows, so that capital costs are accounted for only by interest payments, and not by repayment of principal. Figure 5a-g shows flow diagrams of the computation of physical and economic suitability.

Land utilization types may be defined to have any number of outputs, produced by any number of crops during any number of years. This enables rotations, multiple cropping systems, and intercrops to be analysed. All economic calculations are standardized to units of (currency area⁻¹yr⁻¹) (Fig. 5c). Any currency name and areal unit can be specified for an evaluation. Predicted yields are multiplied by output prices, from a separate table, to determine cash values. Outputs can have negative value, so that, for example, loss of topsoil could be reflected in the economic calculation. If an output price changes then the user can alter a single value, i.e. the output's selling price, and recompute the economic evaluation.

Costs of production may be recurrent or one-time, and may depend on land qualities (Fig. 5d) or not (Fig. 5c). Costs are expressed by listing the number of units of inputs that are required to implement the land use or to overcome a limitation. Changing input prices can be propagated to

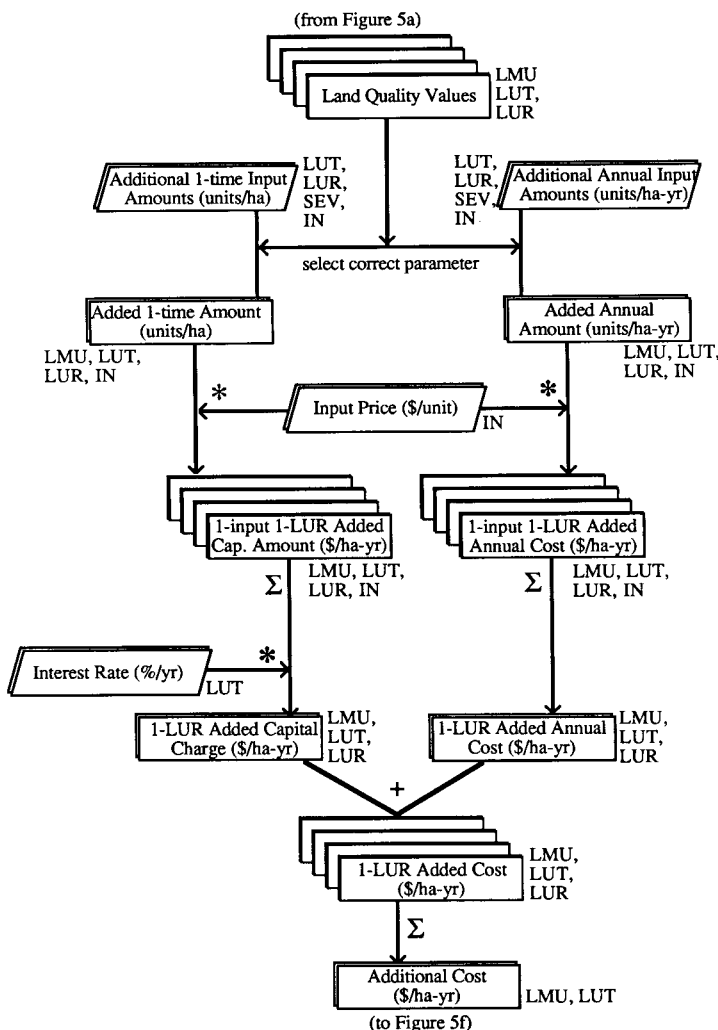


Fig. 5d. Computation of costs that vary with severity levels of land qualities.

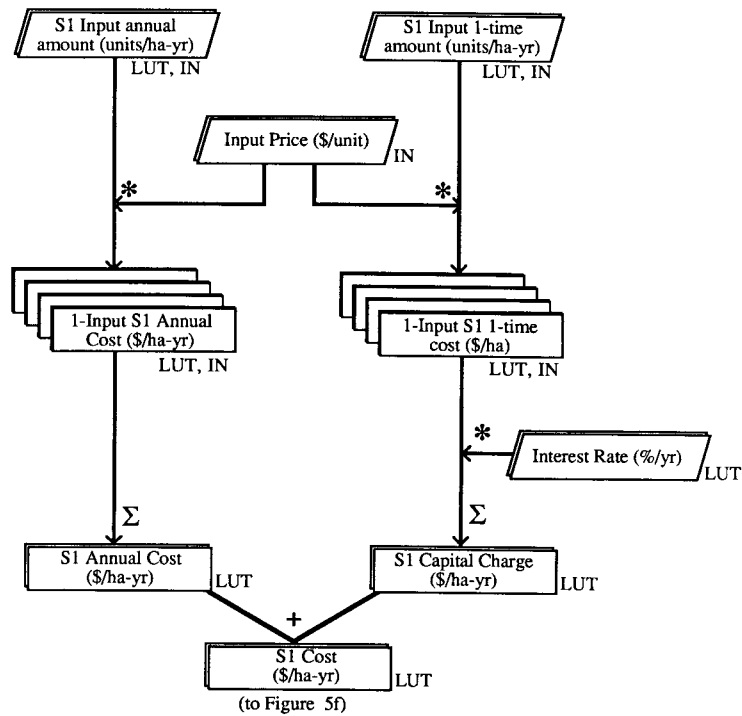


Fig. 5e. Computation of costs that do not vary with severity levels of land qualities.

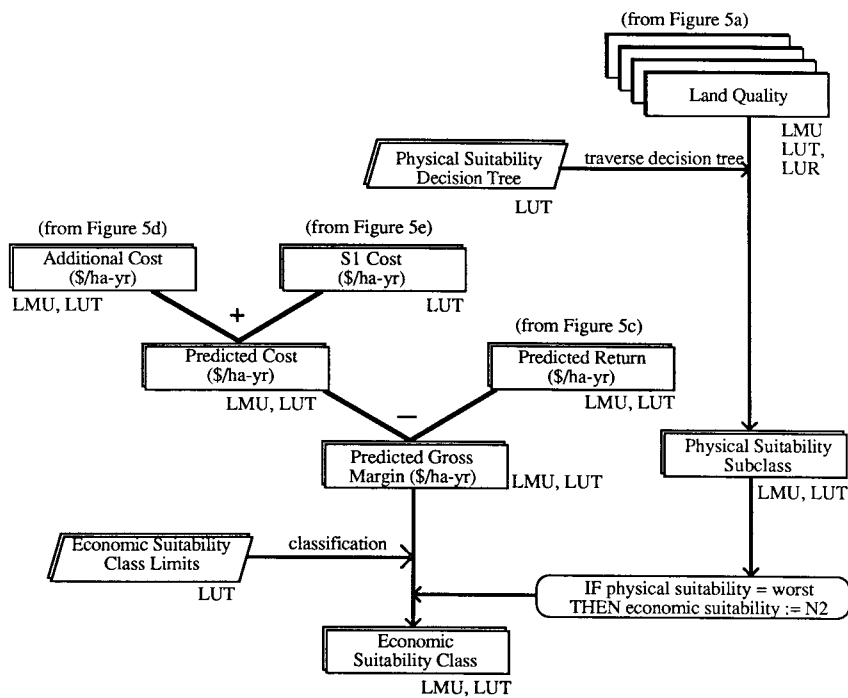


Fig. 5f. Computation of economic and physical suitability classes.

economic suitability values in the same manner as changing output prices.

The key aspect of the economic analysis is that types of land that differ in one or more qualities may differ in their economic suitability. This differentiation is modelled by linking increased limitations for a land quality to increased costs of production, to decreased yields, or to both.

THE ALES DATABASE

Each evaluation model also has its own database. This is a list of the map units whose suitabilities are to be determined according to the model. A *homogeneous* map unit is defined by a list of data values for the various land characteristics used in the model. A *compound* map unit is defined

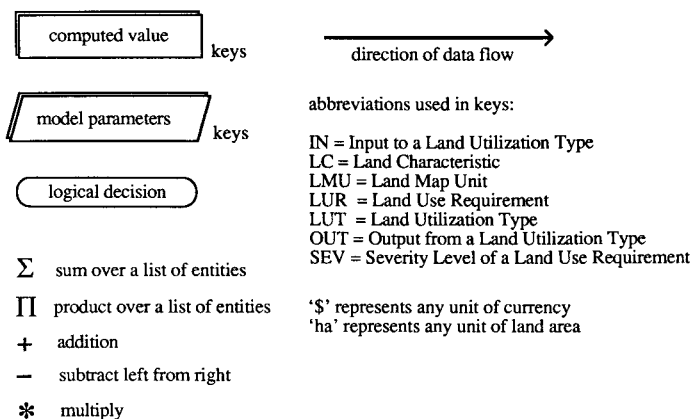


Fig. 5g. ALES computation flow diagram legend.

by a list of its homogeneous constituents, along with the proportion of each constituent in the compound unit. The economic suitability of a compound unit is computed as a linear combination of the economic suitabilities of its constituents.

Land characteristic values may be multiple-valued, each value having an associated linear probability. These probabilities are propagated through the computation, possibly resulting in a multiple-valued evaluation result. In this way a simple probabilistic analysis can be performed from data that vary in time, such as rainfall.

Data may be entered from the keyboard or read from a disc file in a simple interchange format that can be produced by most database programs and spreadsheets. Sets of data

can be written to disc in this same format in order to transfer them to other databases.

ALES and other data processing systems

ALES does not attempt to perform every function that might be needed in a land evaluation project. Figure 6 shows the interfaces between the program and other components in a complete system; these links are explained in the following paragraphs.

The program does not have any geographic referencing. As a result its models cannot easily take into account proximity (e.g. distance to markets) or adjacency (e.g. a LUT requires two kinds of land in a specified pattern). Its models

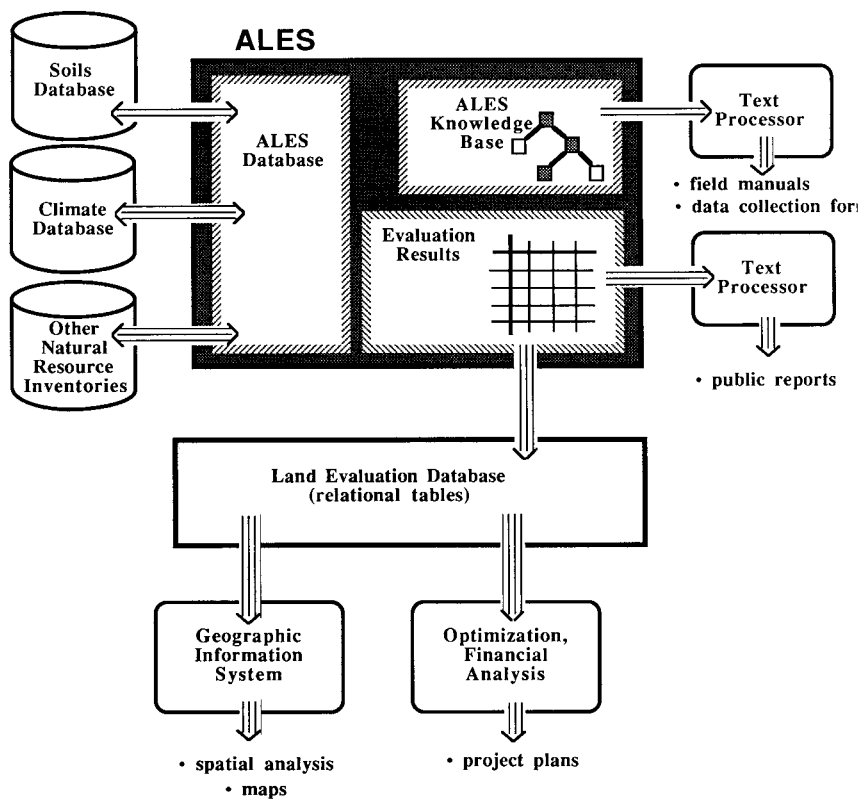


Fig. 6. ALES and other data processing systems.

can make statements only about *map units*, i.e. sets of delineations on a map, and it assumes that the properties of all delineations with the same name are identical, within the precision of the map unit description. This limitation can be overcome by linking it to a geographic information system (Burrough, 1986). Evaluation results, either projected economic returns, crop yields, physical suitability, or individual land qualities, can be written from the program to a relational database. Each map unit's evaluation is a separate record in the relational table. The geographic information system can then be programmed to read the table as one or more overlays.

The program can also exchange map unit data with existing soil or climatic databases. As its data requirements are completely determined by the evaluator for each model, any existing data that can be linked to a map unit can be used in its models.

All model components, most notably the decision trees and annotations, can be written to a standard text file. These can then be edited in text processors to produce handbooks or other instructional materials.

THE ALES PROGRAM

ALES was designed to be used interactively. Context-sensitive 'help' screens are always available at the press of a key. The user selects items and performs actions on them with function keys and menus; there are no commands to enter. The entire context of active windows is usually visible on the screen, and different types of windows can be displayed in different colours on a colour monitor.

Program text, such as prompts, explanations, and menus, are displayed in English by default. However, the user may elect to see any text that has been translated in another language. At present much of the text is being translated to Spanish and French. Because the texts are stored in a database, not in the program itself, translation to other languages does not require any reprogramming.

The program was written in the MUMPS language and database system (Lewkowicz, 1988; MUMPS Users' Group, 1985). It runs on any microcomputer that uses or can emulate the MS-DOS operating system, such as the IBM-PC and its successors. This choice has made the program very widely available. A hard disc is required.

ALES is supplied on a set of diskettes with a procedure to install itself on the user's computer. A comprehensive User's Manual (Rossiter & Van Wambeke, 1989) comes with the program. Four tutorials are included, which lead the novice step-by-step from simple evaluations to the use of the program's more powerful features. The program, manual, and tutorials are available for a fee from the *ALES* Project, Department of Agronomy, Cornell University, Ithaca, NY 14853, USA.

ALES MODELS

The ultimate importance of the program lies in the evalua-

tion models that can be built with it. We expect that in the next few years many governmental agencies, private consultants and educational institutions around the world will use it to produce evaluation models of great utility. Models can be written to a diskette file, from which they can be loaded into other systems. This allows users to exchange models.

At Cornell I constructed several models to illustrate the utility and applicability of the program, using published information as the basis for the decision procedures (Rossiter, 1988). These include the Fertility Capability Classification (Sanchez *et al.*, 1982), several crop requirement tables taken from Sys (1985), and the Zimbabwean land capability classification (Hudson 1981). Implementing these decision-based procedures as models was remarkably easy.

DeRoller (1989) has evaluated the cultivation of selected crops on smallholdings in Guatemala. The decision trees from this model have been collated as a printed manual for field workers who may not have access to a computer. The manual was prepared with a standard word processor after the decision trees has been written to a text file on disc.

An example: land uses for New York State dairy farms

I also constructed a land evaluation of a typical county in New York State for common field cropping systems as an illustrative example. I used as my sources of knowledge standard compendia of local agronomic information (Cornell, 1987, 1989) and estimated farm budgets (Snyder & Lazarus, 1987), a recent soil survey report of the area (Hutton, 1971) and general climate maps (Cornell, 1987). A closer look at this evaluation may help illustrate the nature of the models. The land use types, land use requirements, and land characteristics used in this evaluation need have no relation whatsoever to those that would be used in land evaluations constructed by other evaluators for other areas.

The first step in any land evaluation is defining its objectives, context (area of applicability), and assumptions. In this case the objective was to create a useful and somewhat realistic tutorial, illustrating the expressive power of the program and the use of actual, easily available data. The context is fully mechanized dairy or cash grain farms in central New York State, where farmers are very skilled, where the best current practices are used, and where there is easy access to credit.

The set of land utilization types should ideally be all the reasonable cropping systems and management options. In this example only three illustrative LUTs were defined. The first is continuous maize with conventional practices and without any land improvements (code 'ccc'). The second is similar to the first, but with a permanent land improvement, namely, subsurface drainage (code 'ccc-d'). The third is permanent grass-legume pasture (*Phleum pratense* and *Lotus corniculatus*, code 'tpp').

LUTs are defined within *ALES* by their land use requirements, i.e. the conditions of land that make it more or less

Land Use Requirement	Land Utilization Type		
	ccc	ccc-d	tpp
c (temperature regime)	1	2	2
dr (suitability for subsurface drainage)	x	3	x
e (erosion hazard)	1,2,3	1,2,3	x
f (flood hazard)	1	1	x
fr (frost hazard)	x	x	2
lr (lime requirement)	3	3	3
m (moisture availability)	2	2	2
me (conditions for mechanization)	1	1	1
nr (nutrient requirement)	3	3	3
pl (planting conditions)	2	x	x

legend: 1 - affects physical suitability
 2 - affects yields
 3 - affects costs
 x - not used

Fig. 7. Land utilization types and land use requirements for Tutorial 2.

suitable for the land uses. Figure 7 shows the LURs for the three LUTs. LURs were selected that can either (1) make the land physically unsuitable for the use, (2) reduce yields, or (3) increase costs of production.

An example of a LUR that affects physical suitability is flood hazard for maize (code 'f'). Maize will not tolerate any standing water, so that any significant flood hazard makes production impossible. Hence there are only two levels of the corresponding land quality: no significant flood hazard (less than 1 in 10 years) and otherwise.

An example of a LUR that can reduce yields is moisture availability (code 'm'). Supplemental irrigation is not specified for these LUTs, so that any moisture deficit is reflected in reduced yield. The corresponding LQ is measured in four classes: adequate, slight stress, moderate stress and severe stress.

An example of a LUR that can increase the cost of production is nutrient requirement (code 'nr'). Land with different levels of nutrient supply and retention capacity must be differentially fertilized, but with sufficient expense any land can be brought to a state of optimum chemical fertility, so that yields do not suffer on this account.

A single LUR can have more than one role. For example, erosion hazard for maize (code 'e') has three. The corresponding LQ is rated in four classes: none, slight, moderate, and severe. Land with severe hazard is physically unsuitable for maize under conventional practices. Land with slight or moderate hazard requires contour planting, resulting in increased costs. Land with a moderate hazard also yields less, because 10% of it must be set aside as grass strips.

The LQs that affect yield must be combined to predict the yield. The most general way to combine them is with a *proportional yield decision tree*. Figure 8 shows such a tree for LUT 'ccc', for maize grain. Two LQs, planting conditions and moisture availability, interact to predict a propor-

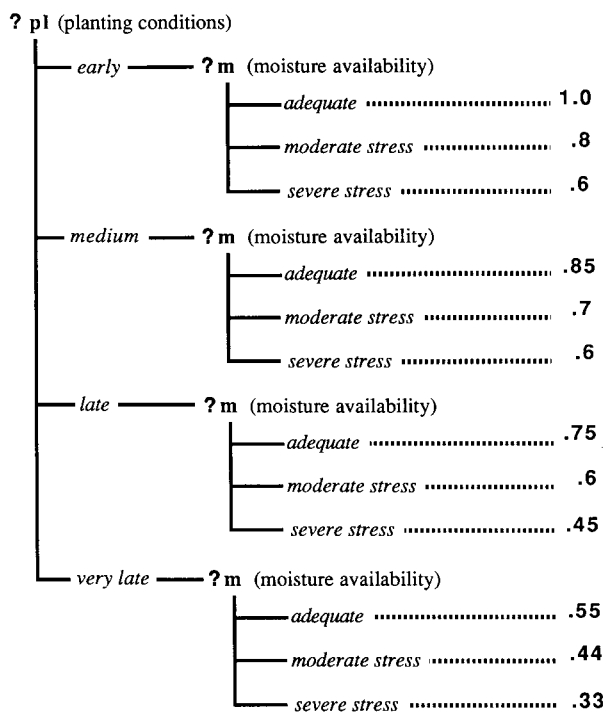


Fig. 8. Proportional yield decision tree for maize grain, in land utilization type 'ccc' of Tutorial 2.

tional yield. Clearly, construction of this tree for many factors would be time-consuming. Therefore ALES provides two ways to deal with LQs that do not interact with others in affecting final yield: *limiting* and *proportional* yield factors.

Some LQs may act according to the agricultural law of the minimum, with no interaction with other factors; an example here is temperature regime. In New York there is a well-established correlation between heat units, measured in growing degree-days, and potential yield of maize (Cornell, 1987), so that it is sufficiently accurate to establish a yield ceiling, i.e. a limiting yield factor, for each of the four levels of the corresponding LQ.

Other LQs may act multiplicatively, with no interaction with other factors; an example here is moderate erosion hazard under conventional maize culture. In this case only 90% of the land may be planted, so that the yield will be 90% of whatever yield is predicted by the yield tree and limiting yield factors.

In addition, land that has severe physical limitations is assumed to have zero yield, so that LQs that influence only physical suitability need not be included in the decision tree. An example here is flood hazard.

Figure 9 shows the land characteristics that were easily available from the two sources. Decision trees had to be constructed so that the program could infer land quality ratings from subsets of these land characteristics. Figure 10 shows a tree for erosion hazard. This land quality can be assessed from a combination of slope, permeability of the upper subsoil, and surface soil texture. Many factors other

Code	Name (Unit of measurement)	No. of classes	
cft-A	type of coarse fragments - surface soil	9	sed in the leaves of the tree). An example here is rainfall intensity. Other factors are constant within the LUT, so do not appear in the decision tree, since a separate tree is built for each LUT. An example here is a cultural-practices factor: it is a constant value for conventionally-tilled maize, and a (different) constant value for, say, no-till maize. Neither LUT's decision tree for erosion hazard would include a cultural-practices factor explicitly; instead, the model builder would hold this factor constant within each tree when determining the land quality values.
cfv-A	volume of coarse fragments - surface soil (%)	4	
dbr	depth to bedrock or root-impenetrable layer (cm)	3	
dc	drainage class	6	
erode	previous erosion	4	
ffs	frost-free season (days)	5	
flood	frequency of flooding (years in 10 with flood)	4	
fpsc	family particle size class from Soil Taxonomy	11	
gdd50	heat units (growing degree days, base 50°F)	7	
perm	permeability of upper subsoil	7	
pm	type of parent material	5	
pm-l	amount of limestone in parent material	3	
pptMS	May-September precipitation (mm)	4	
react	reaction of upper subsoil (pH)	10	
slope	slope of land (%)	5	
text-A	USDA texture class of topsoil	12	
text-B	USDA texture class of subsoil	12	

Fig. 9. Land characteristics used in Tutorial 2 evaluation models.

than these three influence erosion. However, in a restricted area many of these are constants, not variables, and thus need not be considered as branches in the tree (although they do figure into the model builder's decisions as expres-

The New York model: data and results

Once the model was completed a database of some representative map units from the Cayuga County soil survey (Hutton, 1971) was created. Data for soil-related land characteristics were then determined from the descriptions in the soil survey, and values for climatic land characteristics were determined from general climate maps of New York State (Cornell, 1987). These values were entered into the database. Any soil map unit that had delineations with different values of a climatic land characteristic was split into several land map units, each with a single class value of the climatic characteristic. Because both the climate and the distribution of soils are closely tied to topography in this area of New York, only a few soil map units had to be split. Figure 11 shows a representative set of 12 map units,

» <u>slope of land</u>	
-- [A - nearly level]	— 1 (NONE)
-- [B - gently sloping]	» <u>permeability of the upper subsoil</u>
-- --[vs - very slow]	— 4 (SEVERE)
-- --[s, ms - slow to moderately slow]	» <u>USDA surface soil texture</u>
-- -- --[s, ls, sl]	— 2 (SLIGHT)
-- -- --[l, sil, scl, cl, sicl, sic, sc, c]	— 3 (MODERATE)
-- -- --[si]	— 4 (SEVERE)
-- --[m, mr, r, vr - moderate to very rapid]	» <u>USDA surface soil texture</u>
-- -- --[s, ls, sl, scl, cl, sc]	— 1 (NONE)
-- -- --[l, sil, sicl, sic, c]	— 2 (SLIGHT)
-- -- --[si]	— 3 (MODERATE)
-- [C ∅ng]	» <u>permeability of the upper subsoil</u>
-- --[vs, s - very slow to slow]	— 4 (SEVERE)
-- --[ms, m - moderately slow to moderate]	» <u>USDA surface soil texture</u>
-- -- --[s, ls, sl, scl, cl, sc]	— 2 (SLIGHT)
-- -- --[l, sil, sicl, sic, c]	— 3 (MODERATE)
-- -- --[si]	— 4 (SEVERE)
-- --[mr, r, vr - moderately to very rapid]	» <u>USDA surface soil texture</u>
-- -- --[s, ls, sl, scl, cl, sc]	— 1 (NONE)
-- -- --[l, sil, sicl, sic, c, si]	— 2 (SLIGHT)
-- [D - moderately steep]	» <u>permeability of the upper subsoil</u>
-- --[vs, s - very slow to slow]	— 4 (SEVERE)
-- --[ms, m - moderately slow to moderate]	» <u>USDA surface soil texture</u>
-- -- --[s, ls, sl, scl, cl, sc]	— 3 (MODERATE)
-- -- --[l, sil, sicl, sic, c, si]	— 4 (SEVERE)
-- --[mr, r, vr - moderately to very rapid]	» <u>USDA surface soil texture</u>
-- -- --[s, ls, sl, scl, cl, sc]	— 2 (SLIGHT)
-- -- --[l, sil, sicl, sic, c]	— 3 (MODERATE)
-- -- --[si]	— 4 (SEVERE)
-- [E, F - steep to very steep]	— 1 (NONE)

Discriminant entities are introduced by '»' and underlined.
Values of the entities are [boxed].
The level in the tree is indicated by the leader characters, '--'.
Result values are introduced by '—' and written in SMALL CAPITALS.

Fig. 10. Decision tree to determine land quality ratings for erosion hazard in land utilization type 'ccc' Tutorial 2.

Map Unit Code	Map Unit Name
A3*	Honeoye-Lima association
CeB	Cazenovia silt loam, 2-8%
HnB	Honeoye silt loam, 2-8%
HnC3	Honeoye silt loam, 8-14%, eroded
HsD	Honeoye & Lansing gravelly silt loams, 14-20%
KlA	Kendaia & Lyons silt loams, 0-3%
LtA	Lima silt loam, 0-3%
LtB	Lima silt loam, 3-8%
OtF	Ontario, Honeoye, & Lansing soils, 35-50%
OvA	Ovid silt loam, 0-2%
OvB	Ovid silt loam, 2-6%
PgB	Palmyra gravelly loam, 3-8%

* - consists of 4% CeB, 30% HnB, 6% HnC3, 4% HsD, 8% KlA, 14% LtA, 8% LtB, 4% OtF, 4% OvA, 4% OvB, 14% PgB.

Fig. 11. Map units to be evaluated in Tutorial 2.

Land Characteristic*	value
cft-A	n (none)
cfv-A	n (not enough to be named) [0-15%]
dbr	d (deep) [100-1000m]
dc	wd (well drained)
erode	m (moderately eroded)
ffs	150-170 days
flood	n (never)
gdd50	2.2 - 2.4K growing degree days
perm	m (moderate)
pm	till (till, residuum, or colluvium)
pm-1	d (dominant)
pptMS	400-450mm
react	mial (mildly alkaline) [pH 7.3 - 7.8]
slope	B (gently sloping) [3-8%]
text-A	sil (silt loam)
text-B	l (loam)

*see Fig. 9 for translation of land characteristic codes

Fig. 12. Land data values for map unit 'HnB'.

including one association composed of the other 11 units. Figure 12 shows the data for one of the homogeneous map units.

The suitability of each of the twelve map units was then computed for each of the three land utilization types. Figure 13 shows the resulting evaluation matrix. Figure 13a shows the physical suitability subclasses. The physical evaluation was used only to eliminate completely unsuitable land from further consideration. Hence only two physical suitability classes were assigned: 1 (equivalent to FAO classes S1, S2, S3, and N1) and 2 (equivalent to FAO class N2). The subclass suffixes indicate the nature of the limitation. For example, map unit 'OtF' is in physical suitability subclass

Land Mapping Unit	Land Utilization Types		
	ccc	ccc-d	tpp
A3	1 & 2 e & 2e/me	1 & 2e & 2e/me	1 & 2 me
CeB	1	1	1
HnB	1	1	1
HnC	1	1	1
HsD	2e	2e	1
KlA	1	1	1
LtA	1	1	1
LtB	1	1	1
OtF	2e/me	2e/me	2me
OvA	1	1	1
OvB	1	1	1
PgB	1	1	1

Fig. 13a. Evaluation results: physical suitability subclasses.

Land Mapping Unit	Land Utilization Types		
	ccc	ccc-d	tpp
A3	105	92	92
CeB	76	68	95
HnB	155	118	109
HnC	82	42	109
HsD	0	0	105
KlA	71	122	33
LtA	128	122	109
LtB	123	118	109
OtF	0	0	0
OvA	44	69	69
OvB	40	64	70
PgB	100	69	90

Fig. 13b. Evaluation results: gross margin [US \$ acre⁻¹yr⁻¹].

Land Mapping Unit	Land Utilization Types		
	ccc	ccc-d	tpp
A3	102	115	n/a
CeB	95	112	n/a
HnB	128	133	n/a
HnC	97	101	n/a
HsD	0	0	n/a
KlA	90	133	n/a
LtA	115	133	n/a
LtB	115	133	n/a
OtF	0	0	n/a
OvA	81	112	n/a
OvB	81	112	n/a
PgB	108	112	n/a

Fig. 13c. Evaluation results: yield for maize grain [bushels acre⁻¹].

'2e/me' for maize because it has a severe erosion hazard ('e') and severe limitations to mechanization ('me').

Figure 13b shows the predicted gross margins. The column entry (LUT) with the greatest value in each row (map unit) corresponds to the best use, in terms of gross margin, for each land area. This Figure also shows the cost-effectiveness of related LUTs. For example, land drainage is cost-effective on map unit 'KlA' (gross margin increases by US \$51 when going from LUT 'ccc' to LUT 'ccc-d'), but not on map unit 'HnB' (gross margin decreases by US \$37). This is because 'KlA' consists of poorly-drained soils that can be worked much earlier in the growing season if drained, whereas 'HnB' is a medium-textured, freely-draining soil which already allows early planting.

The row entry (map unit) with the greatest value in each column (LUT) of Figure 13b corresponds to the best land areas, in terms of gross margin, for each of the LUTs. For example, the column for LUT 'tpp' shows that the most remunerative areas for permanent pasture are map units 'HnB', 'HnC3', 'HsD', 'LtA', and 'LtB'. This provides a good example of land's being rated for specific uses: land in map unit 'HsD' is totally unsuitable for maize because it is steep and liable to erode, yet it is highly remunerative for permanent pasture, under which erosion is negligible.

Figure 13c shows the predicted yields of maize. Maize is not an output of the pasture LUT, hence this column is marked 'not applicable'. Drainage of map unit 'LtA' increases the predicted yield of maize by about 17%. However, Figure 13b shows that the cost of installing the drainage causes the gross margin on this map unit to decrease by about 4%.

No field investigations were conducted to validate this model. However, its predictions and rankings of map units and land uses were consistent with experience and other rating systems. For example, the predicted yields were very close to the estimated yields for each map unit presented as an interpretive table in the county soil survey (Hutton, 1971). The results, and perhaps more importantly the reasoning embodied in the decision trees, appeared reasonable to conservation officers and extension workers familiar with the area.

CONCLUSION

ALES has proved to be a significant advance in land evaluation. Among its most important innovations are: (1) it is a microcomputer realization of the FAO's framework for land evaluation, suitably enhanced with a definite method of economic analysis; (2) it provides the land evaluator at project level with a highly interactive environment in which to build and refine evaluation models; (3) it uses decision trees to represent the model builder's reasoning explicitly, allowing classified data to be used effectively in decision procedures. The semi-quantitative or even qualitative approach may prove to be especially effective in practice.

ACKNOWLEDGEMENTS

I thank Dr Armand Van Wambeke for conceiving the *ALES* project and supporting it both financially and intellectually. I also thank Mr Michael Tolomeo, Ms Christine DeRoller and Dr Terence Forbes of the International Soils Group at Cornell for their helpful comments. Project Ezra, a cooperative project between Cornell University and the IBM Corporation's Advanced Education Projects programme, provided the IBM PC/AT system on which the program was written.

REFERENCES

- BOLMA, J. & VAN LANEN, H.A.J. 1987. Transfer functions and threshold values: from soil characteristics to land qualities. In: *Quantified land evaluation procedures: proceedings of the international workshop on quantified land evaluation procedures, held in Washington, DC 27 April-2 May 1986*, ITC Publication 6 (eds K.J. Beck, P.A. Burrough & D.E. McCormack) pp. 106-110. International Institute for Aerospace Survey and Earth Sciences, Enschede.
- BURROUGH, P.A. 1986. *Principles of geographical information systems for land resources assessment*. Oxford University Press, Oxford.
- BURROUGH, P.A. 1989. Matching spatial databases and quantitative models in land resource assessment. *Soil Use and Management* 5, 3-8.
- CORNELL UNIVERSITY 1987. *Cornell Field Crops Handbook, 2nd edn*. New York State College of Agriculture and Life Sciences, Ithaca.
- CORNELL UNIVERSITY 1989. *1989 Cornell Recommends for field crops*. New York State College of Agriculture and Life Sciences, Ithaca.
- DENT, D. & YOUNG, A. 1981. *Soil survey and land evaluation*. George Allen & Unwin, London.
- DEROLLER, C. 1989. *Farm scale land evaluation models for land use recommendations in the Guatemalan highlands*. MS Thesis, Cornell University.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS 1976. *A framework for land evaluation*, Soils Bulletin 32. FAO, Rome.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS 1983. *Guidelines: land evaluation for rainfed agriculture*, Soils Bulletin 52. FAO, Rome.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS 1984. *Land evaluation for forestry*, Forestry Paper 48. FAO, Rome.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS 1985. *Guidelines: land evaluation for irrigated agriculture*, Soils Bulletin 55. FAO, Rome.
- HUDSON, N. 1981. *Soil conservation, 2nd edn*. Cornell University Press, Ithaca.
- HUTTON, F.Z. 1971. *Soil survey of Cayuga County, New York*. US Government Printing Office, Washington.
- KLINGEBIEL, A.A. & MONTGOMERY, P.H. 1961. *Land capability classification*. Agriculture Handbook 210. US Government Printing Office, Washington.
- LEWKOWICZ, J. 1988. *The Complete MUMPS*. Prentice-Hall, Englewood Cliffs.
- MUMPS USERS' GROUP 1985. *ANS MUMPS programmers' reference manual*. MUMPS Users' Group, College Park, Maryland.
- O'KEEFE, R., BALCI, O. & SMITH, E.P. 1987. Validating expert system performance. *IEEE Expert* 2, 81-90.
- PURNELL, M. 1987. Quantitative land evaluation in some FAO programs. In: *Quantified land evaluation procedures: proceedings of the international workshop on quantified land evaluation procedures, held in Washington, DC 27 April-2 May 1986*, ITC Publication 6 (eds K.J. Beck, P.A. Burrough & D.E. McCormack) pp. 7-15. International Institute for Aerospace Survey and Earth Sciences, Enschede.
- RIQUIER, J. 1974. *A summary of parametric methods of soil and land evaluation*. In: *Approaches to Land Classification*, Food and Agriculture Organization of the United Nations, Soils Bulletin 22, pp. 47-53. FAO, Rome.
- ROSSITER, D.G. 1988. *The Automated Land Evaluation System: a micro-computer program to assist in land evaluation*, Ph.D. dissertation, Cornell University. University Microfilms, Ann Arbor.
- ROSSITER, D.G. & VAN WAMBEKE, A.R. 1989. *ALES Version 2.2 User's Manual*. Department of Agronomy, Cornell University, Ithaca.
- SANCHEZ, P.A., COUTO, W. & BUOI, S.W. 1982. The fertility capability soil classification system: interpretation, applicability, and modification. *Geoderma* 27, 283-309.
- SIDERIUS, W. (ed.) 1984. *Proceedings of the workshop on land evaluation for extensive grazing (LEEG)*. International Institute for Land Reclamation and Improvement 36, ILRI, Wageningen.
- SIDERIUS, W. (ed.) 1986. *Land evaluation for land-use planning and conservation in sloping areas*. International Institute for Land Reclamation and Improvement 40, ILRI, Wageningen.
- SNYDER, D.P. & LAZARUS, W.F. 1987. *Field crop enterprise budgets: 1987 projections and grower worksheets, New York State*, Mimeo A.E. Res 87-8. Department of Agricultural Economics, Cornell University, Ithaca.
- SYS, C. 1985. *Land evaluation* (Parts I, II, III, and Appendices), Publications Agricoles 7, Brussels: Administration Générale de la Coopération au Développement.
- WATERMAN, D.A. 1985. *A guide to expert systems*. Addison-Wesley, Reading, Massachusetts.
- WOOD, S.R. & DENT, F.J. 1983. *LECS: a land evaluation computer system*, AGOF/INS/78/006/Manual 5 (Methodology) and 6 (User's Manual), Ministry of Agriculture, Government of Indonesia, Bogor.
- YOUNG, A. & GOLDSMITH, P.F. 1977. Soil survey and land evaluation in developing countries: a case study in Malawi. *Geographical Journal* 143, 407-438.