An Operational GIS Expert System for Mapping Forest Soils

Andrew K. Skidmore, Fiona Watford, Paisan Luckananurug, and P. J. Ryan

Abstract

The successful integration of a Bayesian expert system with a commercially available geographic information system (GIS) (Genamap) is described. The package mapped forest soils into five soil landscape classes by utilizing a digital terrain model and vegetation map, as well as knowledge provided by a soil scientist. It is concluded that the map produced by the expert system was as accurate as the map drawn by the soil scientist, within a 95 percent confidence interval. An overall mapping accuracy of 69.8 percent was achieved for the soil maps produced by the expert system, while the conventionally derived map had an accuracy of 73.6 percent.

Introduction

A geographic information system (GIS) is usually viewed as a "tool box" of commands for the input, analysis, storage, retrieval, and display of spatially related data (Tomlin, 1987). Commercial GIS packages, as well as public domain GIS software, have a reasonably standard set of commands to undertake routine operations. For example, conceptually simple data operations such as overlay, buffering, attribute selection, and digitizing are available in virtually all GIS. However, many users are now wishing to integrate human or expert knowledge into their GIS analyses, for more sophisticated processing of their data. Expert systems, also known as knowledge-based systems, have been combined with GIS in a diverse range of applications: e.g.,

- forest fire modeling (Davis *et al.*, 1986),
- scheduling silvicultural practices in forests (Rauscher and Cooney, 1986),
- identification of homogeneous training areas for analysis of remotely sensed imagery (Goodenough *et al.*, 1987),
- forest vegetation mapping (Skidmore, 1989c),
- forest soil mapping (Skidmore *et al.*, 1991),
- digital terrain models (Mackay *et al.*, 1992), and
- local government planning (Davis and McDonald, 1993).

These applications may be characterized as being research orientated studies. The above studies have all indicated that expert systems allow the derivation of dependent GIS layers, and that the dependent GIS layers may not have been successfully derived using other methods. It is the use of expert (human) knowledge which allows the expert system to generate results which are similar to those expected from a human expert (Forsyth, 1984). In general, such studies have not attempted to rigorously quantify the accuracy of the expert system compared with the expert human.

The aim of this study was to prove that expert systems may be incorporated into a GIS for operational use. In order to successfully implement an operational expert system with a GIS, we felt that a convenient interface between the user and the GIS must be developed, as well as prove that the expert system produces a map layer which has an accuracy comparable to the map produced by a human expert. Thus, specific objectives were to

- Construct an intuitive user interface for a commercial GIS, thereby allowing non-specialist users access to the technology;
- Rigorously test the accuracy of the output from the expert system, and statistically compare the expert system map with a conventional map produced by a human expert; and
- Indicate whether the output from the expert system is of an accuracy that would be considered operational.

It should be emphasized that the testing of the expert system was performed in a specific application domain; that is mapping forest soils. Nevertheless, techniques presented here should be applicable to many other GIS applications, if the analyst has sufficient knowledge about the application, and the independent GIS data layers are available for input to the expert system.

What Are Expert Systems?

Expert systems are computer programs which use symbolic knowledge to simulate the behavior of human experts (Stock, 1987), and they are a topical issue in the field of artificial intelligence (AI). However, people working in the field of AI continue to be confused about what AI really is (Schank, 1988). In other words, there are attempts to confer properties (or attributes) to a computer system under the guise of AI, but the practitioners find difficulty in defining these properties! It is generally accepted that an expert system is useful when it reaches the same conclusion as an expert (Weiss and Kulikowski, 1984). But, as Halpern (1987) points out, one cannot define artificial intelligence (and, *ipso facto*, an expert system) in terms of Turing's Test, which Halpern (1987) redefines as "... Can a computer be programmed so as to fool human beings, with some regularity, into thinking that it is another human being?"

In this study, we view expert systems as comprising a knowledge base, an inference engine, and a user interface (Figure 1). The knowledge base contains facts and rules which the program uses to search for a solution to the prob-

A.K. Skidmore and F. Watford are with the Centre for Remote Sensing and GIS, School of Geography, University of New South Wales, P.O. Box 1, Kensington, NSW 2033, Australia.

P. Luckananurug is with the School of Computer Science and Engineering, University of New South Wales, P.O. Box 1, Kensington, NSW 2033, Australia.

P.J. Ryan is with the CSIRO Division of Forestry and Forest Products, PO Box 4008 QVT, Canberra, ACT 2600, Australia.

Photogrammetric Engineering & Remote Sensing, Vol. 62, No. 5, May 1996, pp. 501–511.

0099-1112/96/6205–501$3.00/0

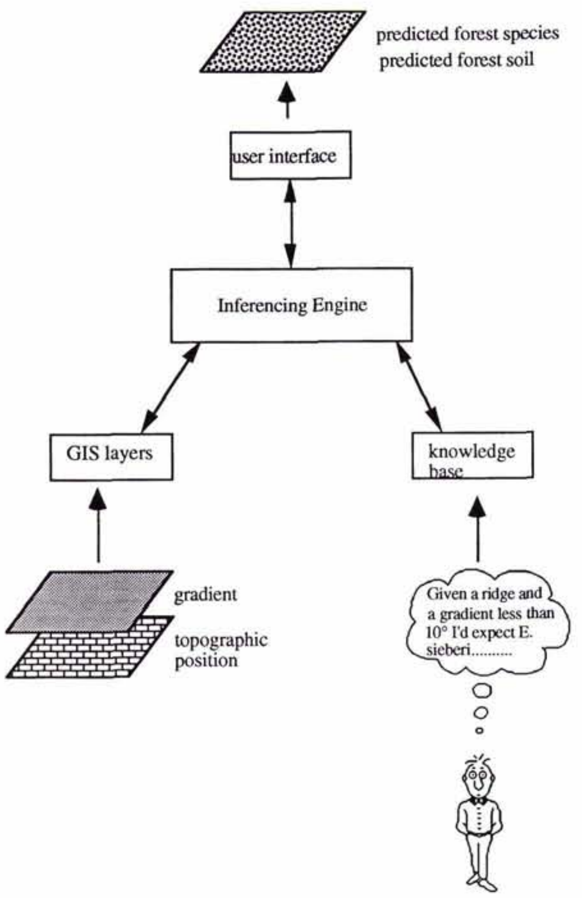

Figure 1. Components of an Expert System.

lem (Stock, 1987); in this case, to find the most likely forest soil occurring at a grid cell. The inference engine uses the knowledge base to infer logically valid conclusions, and to logically justify conclusions at the completion of the program (Davis and McDonald, 1993). In a conventional expert system, the user supplies input about the problem to the inference engine (for example, the user may be prompted by a medical diagnostic expert system to state the temperature of a patient). Implementing a GIS-based expert system requires that the input to the program is derived from the cell (or polygon) attributes of the GIS (Figure 1).

The expert system used in this study is a Bayesian implementation. The basis of the Bayes' inferencing algorithm is that knowledge about the likelihood of a hypothesis occurring, given a piece of evidence, may be thought of as a conditional probability. For example, a user may not be certain whether *Eucalyptus sieberi* always occurs on ridges — it may sometimes occur on midslopes. The knowledge may be expressed as the user being 90 percent certain (i.e., probability = 0.9) that *Eucalyptus sieberi* occurs on ridges. In Figure 1, the knowledge base may be represented as a probability matrix (Skidmore, 1989b), as a modified frame (Skidmore *et al.*, 1992), or by using a mega-rule structure as described in the next section.

Again, referring to Figure 1, GIS data layers are conventional raster grid or overlaid polygon layers (Burrough, 1986). Attributes of the raster cell or polygon are input to the expert system and matched with the information in the knowledge base. The expert system then infers the most likely class at a given cell, using Bayes' Theory to update the probability of the rule that the hypothesis (H_a) occurs at a grid cell location given a piece of evidence (E_b), i.e.,

$$P(H_a \mid E_b) = \frac{P(E_b \mid H_a)\, P(H_a)}{P(E_b)}. \tag{1}$$

$P(E_b \mid H_a)$ is the *a priori* conditional probability estimated by the expert. $P(H_a)$ is the probability for the hypothesis (H_a) that class S_a occurs at location (i,j) and is estimated by the expert. On iterating with the $b = 2, ..., k$ items of evidence from the GIS database, $P(H_a \mid E_b\ ;\ b = 1)$ (i.e., the *a posteriori* probability of H_a given E_b, for $b = 1$) replaces $P(H_a)$ in Equation 1. $P(E_b)$ is the classical marginal probability, and is the probability of the evidence alone, or the probability that any cell has an item of evidence $\{E_b\}$ such as a southerly aspect. Bayes' Theorem provides a formula to calculate $P(E_b)$: i.e.,

$$P(E_b) = \sum_{a=1}^{n} P(E_b \mid H_a)\, P(H_a),$$

thereby allowing $P(E_b)$ to be continually updated at runtime as $P(H_a)$ is updated.

Two methods exist for linking the evidence with the hypotheses. The first is forward chaining, where the inference works forward from the data (evidence) to the hypothesis. This is a data-driven process where, given some evidence, a hypothesis is inferred. The second method is simply the reverse, and is called backwards chaining. The inferencing flows (back) from the hypothesis to the data. In other words, given a hypothesis, the expert system examines how much evidence there is to support the hypothesis. Backwards chaining is obviously a hypothesis-driven process. The expert system developed for this study used forward chaining with a complete enumeration of the data (i.e., a blind search terminated by running out of evidence).

The Expert System Knowledge Base

The knowledge base used in this implementation of the expert system consists of facts and rules. Facts describe fixed properties of the knowledge, while rules are used to deduct new facts from existing facts. In other words, a fact may be thought of as a type of passive knowledge which is inherent in the knowledge base, while the rules are active knowledge which are generated by the expert system.

The expert system rules and facts were generated by an experienced soil scientist (P. Ryan), were determined through a knowledge of soil characteristics and the geomorphic processes which contribute to the formation of the soil, and were also based on the intuition of the expert. The *a priori* probabilities were estimated from the approximate areal extent of each soil landscape unit across the area.

The facts record the relation between soil-landscape classes and the environmental evidence. Each fact is expressed as the probability of finding an item of evidence, given a hypothesis (soil-landscape unit), i.e., $P(E_b \mid H_a)$ in Equation 1. Facts are formalized as follows:

fact(probE2,[10,20,50,30,20]),

where the list of numbers in square brackets is assigned to the variable "probE2." For example, $P(E_2 \mid H_1) = 0.1$, $P(E_2 \mid H_2) = 0.2$, $P(E_2 \mid H_3) = 0.5$, and so on. The knowledge about how to use these facts in the classification process is represented by a set of production rules.

Each production rule comprises a label, condition, and action, with the following format:

LABEL: if **CONDITION**
then **ACTION**

The LABEL: uniquely identifies each rule. If the CONDI-

TABLE 1. PRODUCTION RULE ACTIONS

Actions	Description
X and Y	Perform the action X and the action Y
using X to Y	Save the value of X along with the satisfied condition for later use by procedure Y
do Y	Call the procedure Y
read X from Y	Read the value of X from file Y
enter X	Alloy the reasoning to use the rules in mega-rule X
exit X	Prevent the reasoning from using the rule in mega-rule X
write X to Y	Write the value of X to file Y

TION is satisfied (i.e., true), then the ACTION is performed and the rule is said to have fired. For example, one of the rules is as follows:

rule_62: if **evidence = 2**
then **using probE2 to update_prob**

The CONDITION may contain Boolean operators "AND" and "OR," as well as inequalities (i.e., greater than [>], less than [<], equal to [=]). The ACTION part of the rule dictates the processing that the expert system will undertake. Types of actions that the rule may take are detailed in Table 1.

The rules may be grouped into mega-rules, in which the rules will only be checked if that mega-rule is active. The "enter" and "exit" actions (Table 1) are used to activate and deactivate the mega-rules, respectively. The use of mega-rules allows easier editing of the knowledge base, improves the response time of the system because irrelevant rules are not tested, and organizes knowledge around objects or topics of interest.

Using a conventional forward chaining process, all rules would need to be evaluated for a given pixel. In other words, the number of grid cells would increase the number of computations in a linear manner. For example, if there are 100 by 100 grid cells, ten-thousand passes through the rules would be required. For a set of 100 rules, one-million pattern matching operations would need to be performed.

To speed up the algorithm, the action part of the production rule may create a new fact, which then triggers another rule. For example,

fact (probE21,[70,20,40,50,60])
fact (probE22,[30,60,50,40,40])

rule_61: if evidence = 1
then using probE21 to update_prob and
assign evidence = evidence + 1

rule_62: if evidence = 2
then using probE22 to update_prob and
assign evidence = evidence + 1

The processing cycle may be interpreted as follows. When rule_61 is fired, the "assign" action increases the value of the "evidence" variable by 1. This causes rule_62 to fire. The "using" action in each rule saves the evidence value along with the probability value (that is probE21 and probE22) into an internal array. This array value is used later by a procedure called "update_prob," which calculates the value of each cell according to Equation 1. By encoding the knowledge base in this manner, the set of rules is traversed only once, resulting in the execution time for the expert system being proportional to the number of rules. The actual implementation of this algorithm involves writing the evidence and probability values to an array maintained for each map layer. A class is assigned to each grid cell by pattern matching with the internal array variable. This is a much faster process compared with the rule condition checking mechanism.

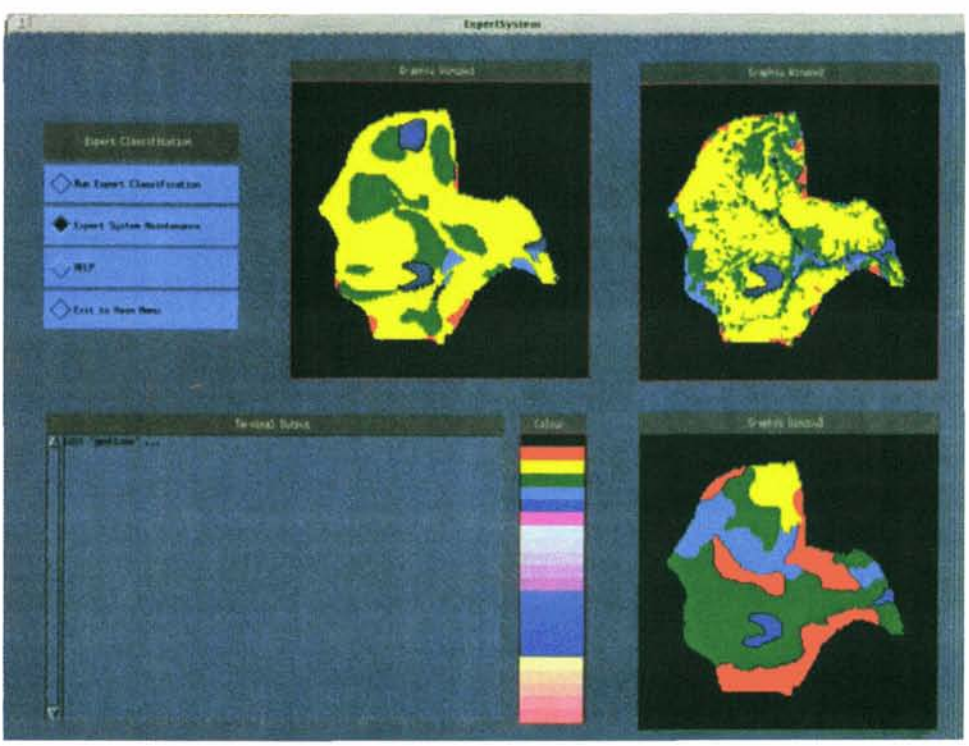

Plate 1. User interface for LCMES.

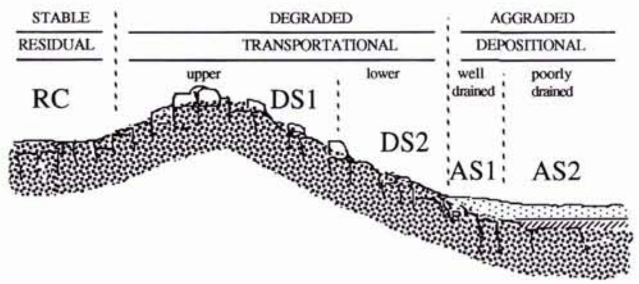

Figure 2. Geomorphological model showing the five soil-landscape units recognized in this study (after Ryan, 1992).

LCMES
The Land Classification and Mapping Expert System (LCMES) is an implementation of the Bayesian expert system initially described by Skidmore (1989b). LCMES consists of four main components: knowledge base, GIS, inference engine, and user interface. The user interacts with the system through a menu, using a mouse control interface (Plate 1). The main features of LCMES are

- the program has been implemented using the Genasys GIS, incorporating the Genius user interface;
- the user can input, examine, and modify the relevant data;
- the user can combine the evidence stored in the GIS layers with the classification process;
- the user can store the classification results produced by the system in the GIS data base;
- the knowledge base may be modified through a text editor; and
- the resulting thematic map can be displayed through a graphic window, or it can be printed.

It is possible to apply the expert system to most GIS problem domains. In the example cited below, the expert system mapped forest soils using terrain parameters derived from a digital elevation model, a forest overstory layer interpreted from aerial photographs, and a knowledge base generated by interviewing a soil scientist. Using the editing features of LCMES, it is possible to modify the knowledge base, while the GIS data layers may be added, deleted, or modified using the GIS. The application of the package is only limited by the imagination of users, availability of the GIS data layers, and knowledge about the variable being modeled.

Testing the Accuracy of the Expert System

Introduction
In order to test the accuracy of the expert system, one soil map was produced by the expert system and one by a soil scientist using conventional methods. In this section, the derivation of the maps is described, and the statistical methods used to assess accuracy are detailed.

Study Area
Two catchments were selected to test the expert system. Both catchments are located 30 km inland on the far south coast of New South Wales (approximately 50 km south west of the town of Eden). The catchments, named Geebung Creek (79 ha) and Peppermint (128 ha), respectively, were chosen because of a well established soil-landscape unit model developed by Ryan (1993) which forms the basis of comparison with the expert system results.

The vegetation is native dry schlerophyll forest, with a tall open structure that has a maximum height of approximately 35 m. A complex mix of approximately eight common eucalypt species intergrade to form the overstory. Both catchments occur on Wallagaraugh Adamellite, which was described by Beams (1980) as a coarse grained pink felsic adamellite/granite. The relatively homogeneous parent material results in soil development being strongly influenced by geomorphic processes related to slope evolution, which produced a series of interrelated soil landscapes (Ryan, 1993).

The study area has a moist mid-latitude climate with an average annual precipitation of 892 mm which falls uniformly through the year. The mean temperatures range from a summer maximum of 27.3°C to a winter minimum of 3.1°C (Ryan, 1993).

Application Domain and the Soil Landscape Model
Skidmore *et al.* (1991) and Ryan (1993) describe a soil landscape model for the two catchments. The classification of the units is based on air-photo interpretation and field visits, and uses the terminology of Parker (1991). The soil landscape model (Figure 2) has three major geomorphic environments:

- residual surfaces on the watersheds and broad hillcrests,
- transportational surfaces associated with steeper slopes and ridges, and
- depositional surfaces associated with the footslopes and valley floors.

These geomorphic zones may be divided into five soil landscape units as described by Skidmore *et al.* (1991) and Ryan (1993): residual crests and interfluves (RC), degraded mid- to upper slopes (DS1), degraded lower slopes (DS2), aggraded well drained slopes (AS1), and aggraded slopes with restricted drainage (AS2).The relative position of these five soil-landscape units is shown in Figure 2.

Data Layers
The independent variables input to the expert system were constructed as a raster database within the Genasys GIS, with a 10-m grid cell size and projected to a UTM coordinate system. The independent variables are represented as the layers or themes at each grid cell in the GIS. The following layers were input to the GIS:

Vegetation. Vegetation typing from aerial photographs and ground visits was completed by an experienced consultant using the Baur (1965) forest classification system for New South Wales. The forest types were digitized from the aerial photographs, and a vector-to-raster conversion was performed. In all, ten forest types were recognized over the area (Plate 2).

Digital Elevation Model. Contours, high points, saddles, and streamlines were digitized from a 1:5000-scale topographic map, and were interpolated to the 10-m cell size using the Splin2H program developed by Hutchinson (1989). The use of this interpolation model is described in detail in Skidmore (1989a) and Skidmore (1990).

Soil Wetness Index. The CSIRO Centre for Catchment Hydrology allowed the use of the program TOPOG initially developed by O'Loughlin (1986). The model produces a map of relative saturation, using the gridded digital elevation model output by Splin2H (Hutchinson, 1989) (see Plate 3). Note that T is the transmissivity and b is the base flow or efflux assumed for the TOPOG model.

Topographic Position. Gridded digital elevation model output by Splin2H (Hutchinson, 1989) was also used as input to a program developed by Skidmore (1990) to classify the topography into five classes: ridge, upper mid-slope, mid-slope, lower mid-slope, and gully (Plate 4).

Slope. The slope data layer was calculated by using the second-order finite difference method implemented by Skidmore

(1989a). The slope classes were reclassed into areas of less than 5°, areas of 6 to 10°, areas of 11 to 20°, and areas of greater than 21° (Plate 5).

The topographic position and soil wetness layers were overlaid (combined) in order to increase the amount of expert knowledge that could be recorded for each soil landscape unit. For example, a gully high in the catchment would be a dry gully, while a gully low in the catchment would be a wet gully. These different types of gullies were known to be associated with different soil types, and so specific expert rules could be generated.

Soil-Landscape Knowledge Base

An example of the knowledge base developed for this application is shown in Appendix A for the five soil landscape unit classes, as well as for the three GIS layers (a combined soil wetness-topographic position layer, a slope gradient layer, and a forest overstory layer). The structure of the knowledge base is as described above in the section entitled The Expert System Knowledge Base.

Soil Sampling and Accuracy Assessment

To test the accuracy of the conventional soil map and the soil map derived by the expert system, soil pit profiles were described across the two catchments. For each soil pit, a field assessment of the soil landscape units was made based on the descriptions of Ryan (1993). It included depth of soil, degree of horizonization, the colors of the horizons, and field texture of the soils.

The sampling design chosen was a stratified random sample, with the strata based on soil landscape units described by Ryan (1993). A stratified random sample has the advantage that sample units are distributed equitably over the area, and that small but important areas are included in the sample (Berry and Baker, 1968; Congalton, 1988).

Hay (1979) concluded that a minimum sample size should be 50. He stated that " ... any sample size of less than 50 will be an unsatisfactory guide to error rates". Congalton (1988) empirically simulated the effect of varying sample size on the estimated population parameters. He found highly variable population estimates, which appeared to depend on sample method as much as number of samples. Generally, simple random sampling gave the best results, requiring between 50 and 100 samples per category in order to approximate sample parameters in a stable manner. For these reasons, we selected a sample size of approximately 50 (actually, 53 samples were finally recorded), with at least ten samples occurring in each of the five soil landscape units.

These above studies by Hay (1979) and Congalton (1988) calculate sample size based on confidence limits or acceptance testing. Thomas and Allcock (1984) described a method for calculating confidence intervals about a mapping accuracy statement, for a sample of a size specified by the user. Their technique is based on using binomial distribution theory. An assumption is that the minimum number of samples should be greater than 50. The 99.9 percent confidence level (CL) for a map's accuracy is calculated as follows:

$$99.9\%\ \mathrm{CL} = (m - 3e_m) - 3(s + 3e_s) \tag{2}$$

where

N = number of samples taken,
p = number of samples that have been correctly classified,
q = $1 - p$,
m = Np,
s = $\sqrt{(Npq)}$,
e_m = $s/\sqrt{N}$, and
e_s = $s/\sqrt{(2N)}$.

To calculate the 99 and 95 percent confidence levels, the 3 in Equation 2 can be replaced with 2.33 and 1.65, respectively. Note that this technique can be used on a class by class basis, or for the whole image. Assuming that 70 percent of the sample points would be assigned to the correct class, different numbers of samples may be substituted into the equation to investigate how the map accuracy changes with different confidence limits. For example, taking 50 samples with a 95 percent confidence limit, the map accuracy is 56 percent. If the sample size is increased to 100, the map accuracy only improves by 6 percent, to 61 percent. Thus, the increase in mapping accuracy, for a given confidence limit, is small, especially considering the huge field effort required.

A grid system of 95-m squares in Geebung Creek catchment, and 125- by 110-m rectangles in the Peppermint catchment had been precision surveyed, as part of an earlier hydrological experiment. A random sample of grid coordinates was calculated within the catchments, and was plotted onto 1:5000-scale topographic maps superimposed with the grid. It is estimated that the points were located within ± 3 metres of the location recorded on the map. As discussed above, the grid resolution for the GIS was 10 metres, so the soil landscape unit class of each sample point could be justifiably assigned to the grid cell over which it occurs.

Cohen (1960) and Bishop *et al.* (1975) defined a measure of overall agreement between image data and the reference (ground truth) data called Kappa or K: i.e.,

$$K = \frac{\theta_1 - \theta_2}{1 - \theta_2}$$

where

$$\theta_1 = \sum_i p_{ii} \text{ and } \theta_2 = \sum_i p_{i+}\, p_{+i}\,.$$

Note that p_{i+} is the sum of the ith row and p_{+i} is the sum of the ith column. p is the simple proportion obtained by dividing the observed counts in the error matrix by the total number of observations N.

They further defined the estimated asymptotic variance of K: i.e.,

$$\sigma_\infty^2[\hat{K}] = \frac{1}{N}\left\{\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2-\theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4-4\theta_2^2)}{(1-\theta_2)^4}\right\}$$

where θ_1 and θ_2 are as above and

$$\theta_3 = \sum_i p_{ii}\,(p_{i+} + p_{+i}) \text{ and } \theta_4 = \sum_i p_{ij}\,(p_{i+} + p_{+j})^2.$$

K ranges in value from 0 (no association, that is, any agreement between the two images equals chance agreement) through to 1 (full association, there is perfect agreement between the two images). K can also be negative, which signifies a less than chance agreement. Rosenfield and Fitzpatrick-Lins (1986) made the point that values for K (expressed as a percentage, i.e., $K \times 100$) are less than the values for total percent correct (or mapping accuracy).

To test for a statistically significant difference between two error matrices, Cohen (1960) proposed using the K values (e.g., K_1 and K_2 representing maps 1 and 2, respectively) and their associated variance by evaluating the normal curve deviate: i.e.,

$$z_t = \frac{K_1 - K_2}{\sqrt{\sigma_{K_1}^2 + \sigma_{K_2}^2}}\,.$$

This test statistic may be applied to paired combinations of error matrices in order to ascertain whether the error ma-

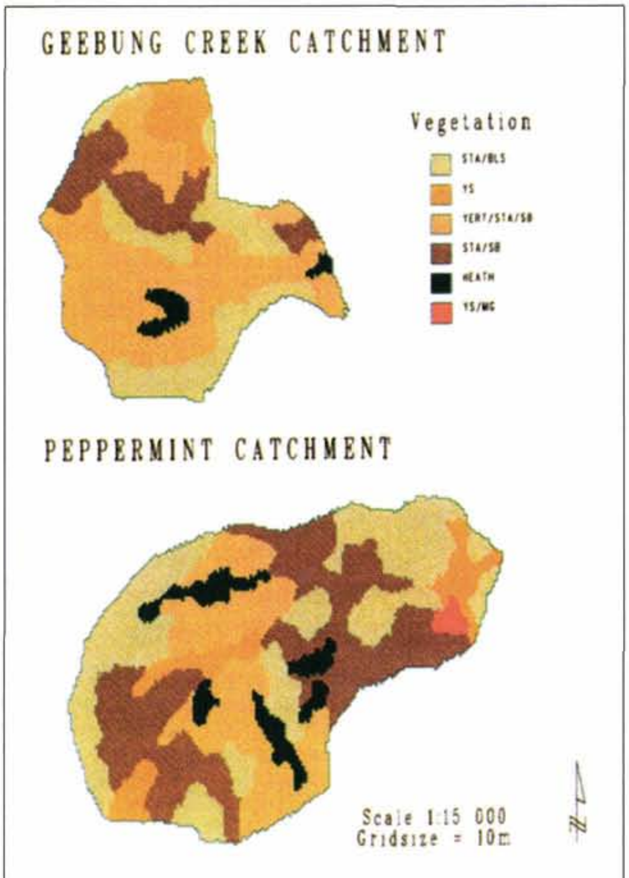

Plate 2. Forest types.

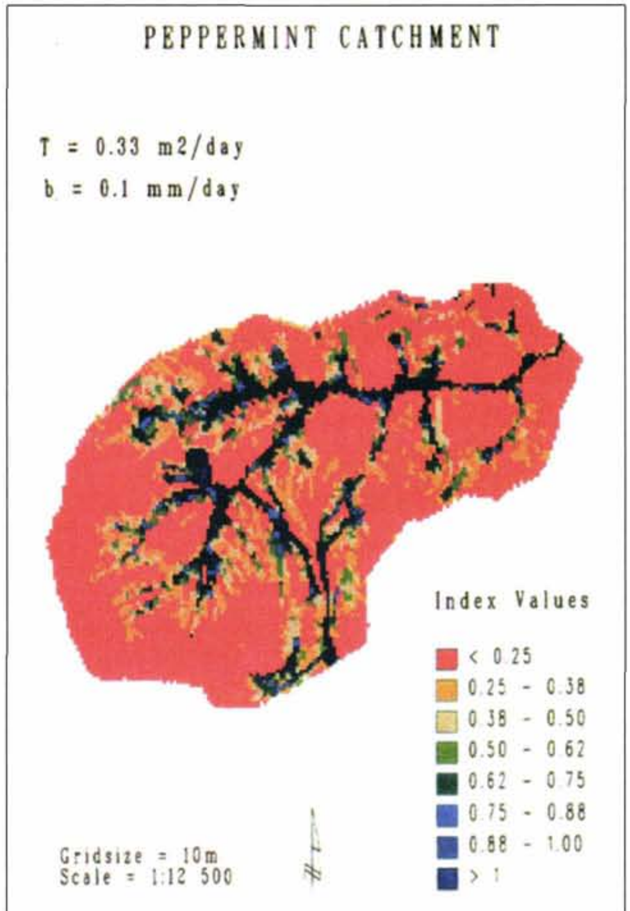

Plate 3. Relative saturation over the Peppermint Catchment.

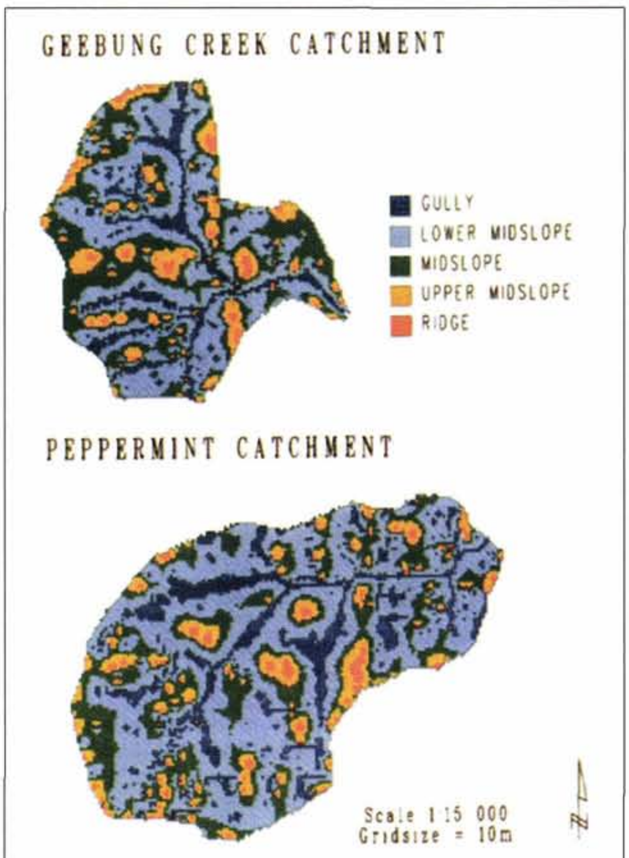

Plate 4. Topographic position.

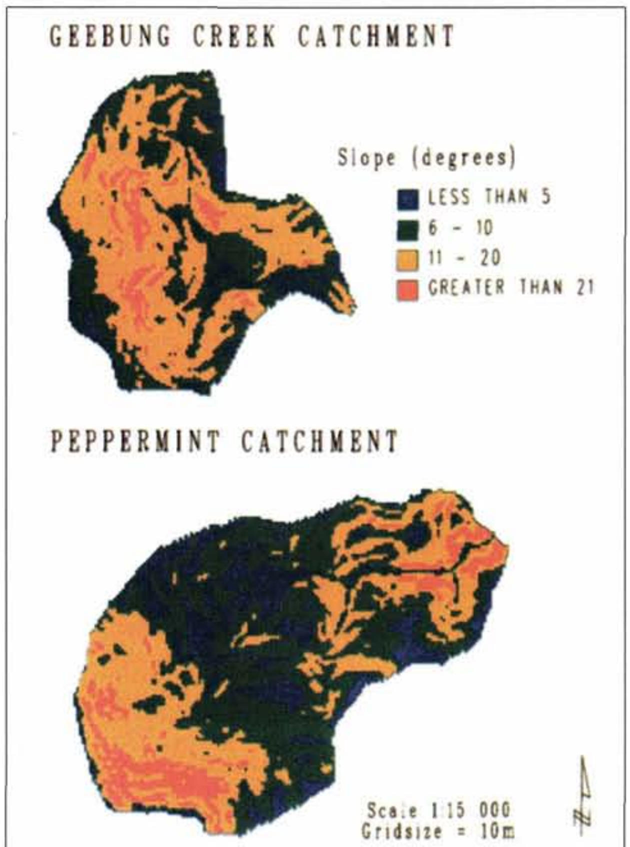

Plate 5. Slope gradient classes.

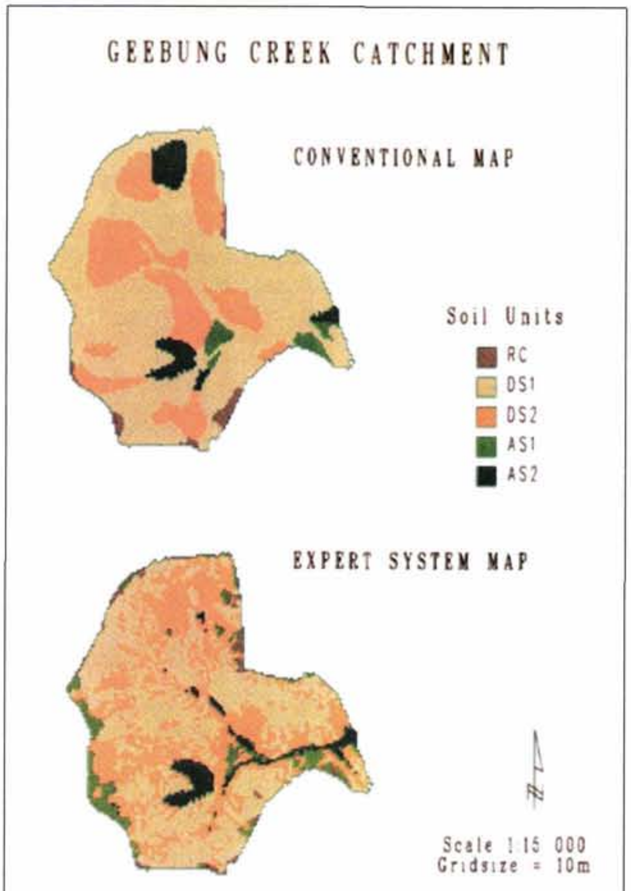

Plate 6. Map of the Geebung Creek catchment produced by the expert system and soil scientist.

trices are significantly different. A null hypothesis can be set up to test whether the K values for the two maps differ: i.e.,

$$H_O: K_1 = K_2$$

versus

$$H_a: K_1 \neq K_2$$

The null hypothesis is rejected using the normal curve deviate statistic (z) for $\alpha = 0.05$ if $z_t > 1.96$ (i.e., $z_{\alpha=0.05} = 1.96$). Note that any other rejection region can be used, e.g., $\alpha = 0.01$ or $\alpha = 0.001$.

Examples of the K statistic applied to remote sensing problems can be found in Congalton *et al.* (1983) and Skidmore (1989b). Rosenfield and Fitzpatrick-Lins (1986) discussed the Cohen (1960) K coefficient as a relation to a family of coefficients which correct for chance agreement between two error matrices (or contingency tables). They commended the Cohen K coefficient statistic because it considers within-class correlation as well as overall image correlation. In other words, all cells in the error matrix are considered (Fung and LeDrew, 1988).

Results

Plates 6 and 7 show the maps of Geebung Creek and Peppermint catchments produced by the expert system and the soil scientist. These figures allow a visual comparison of the two techniques.

Confusion matrices for the two catchments were calculated for both the conventional map derived by the soil scientist, and for the expert system maps (Tables 1 to 4). The confusion matrices for the conventional soil mapping were combined for both catchments (Table 5), as were the confusion matrices for the expert system (Table 6). The overall mapping accuracy is summarized in Table 7. The overall mapping accuracy is the number of correctly mapped sample cells divided by the total number of cells sampled.

The Thomas and Allcock (1984) method was used to determine the accuracy of the maps within 95 percent confidence intervals (Table 8).

From Table 7 and 8, it can be seen that there is little difference in the accuracy of the conventional map and the expert system. Using the Kappa (K) statistic (Cohen, 1960; Congalton *et al.*, 1983), we tested whether there is a significant difference in the accuracy of the two maps. The computed test statistic is $z = 0.44$, which is less than the critical z value of $z = 1.96$ (for a 95 percent confidence interval). Thus, we accept the null hypothesis (H_o) and conclude, with 99 percent confidence, that there is no significant difference between the accuracy of the conventionally derived map and the map produced by the expert system.

Discussion

There are obvious visual differences between maps produced by the expert system and the conventional method, though statistically there is no difference between the two techniques. Why is this? The most likely answer is that the maps are equally accurate (if you believe the statistics) and that the error in the conventional map is introduced by the interpreter. In other words, the map produced by the soil scientist looks accurate because it appears to have large homogeneous classes with smooth edges — in fact, it is in a

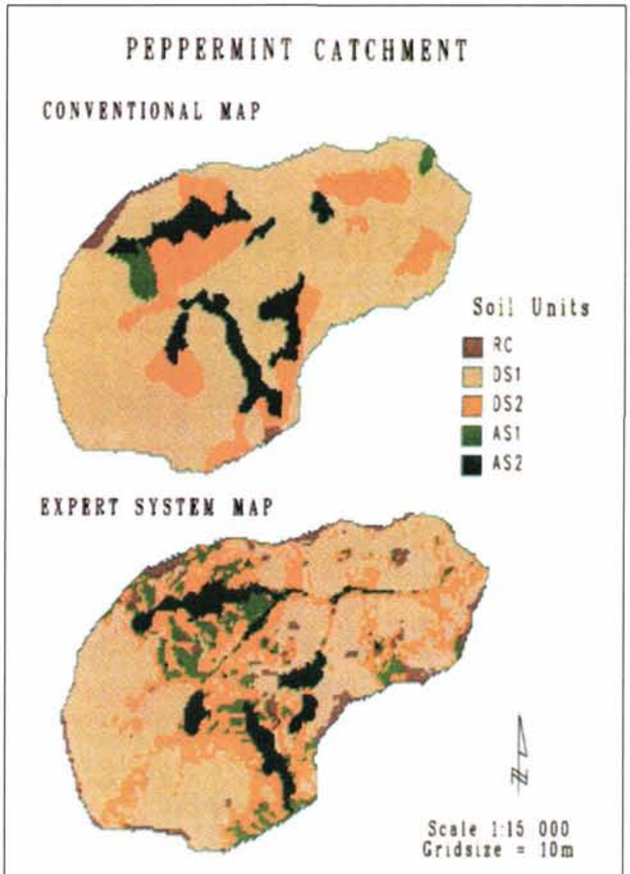

Plate 7. Map of the Peppermint catchment produced by the expert system and soil scientist.

Table 1. Confusion Matrix of Conventional Soil Map, Geebung Creek Catchment

	Sample Points					
Soil Map	RC	DS1	DS2	AS1	AS2	Total
RC	4					4
DS1	2	4				6
DS2		2	3			5
AS1				4		4
AS2			1		3	4
Total	6	6	4	4	3	23

Table 2. Confusion Matrix of Conventional Soil Map, Peppermint Catchment

	Sample Points					
Soil Map	RC	DS1	DS2	AS1	AS2	Total
RC	3					3
DS1	1	4	1	2	1	9
DS2		2	5		1	8
AS1			1	4		5
AS2					5	5
Total	4	6	7	6	7	30

Table 3. Confusion Matrix of the Expert System Soil Map, Geebung Creek Catchment

	Sample Points					
Soil Map	RC	DS1	DS2	AS1	AS2	Total
RC	4			1		5
DS1		4	1			5
DS2		2	3			5
AS1	2			3		5
AS2					3	3
Total	6	6	4	4	3	23

Table 4. Confusion Matrix of the Expert System Soil Map, Peppermint Catchment

	Sample Points					
Soil Map	RC	DS1	DS2	AS1	AS2	Total
RC	2			1		3
DS1		5	1	1		7
DS2		1	4	1	1	7
AS1	2			3		5
AS2			2		6	8
Total	4	6	7	6	7	30

Table 5. Confusion Matrix of Conventional Soil Map, Combined Geebung Creek and Peppermint Catchments

	Sample Points					
Soil Map	RC	DS1	DS2	AS1	AS2	Total
RC	7					7
DS1	3	8	1	2	1	15
DS2		4	8		2	13
AS1			1	8		9
AS2			1		8	9
Total	10	12	11	10	10	53

Table 6. Confusion Matrix of the Expert System Soil Map, Combined Geebung Creek and Peppermint Catchments

	Sample Points					
Soil Map	RC	DS1	DS2	AS1	AS2	Total
RC	6			2		8
DS1		9	2	1	1	13
DS2		3	7	1		11
AS1	4		2	6		12
AS2					9	9
Total	10	12	11	10	10	53

Table 7. The Overall Mapping Accuracy (Percent) of the Conventional and Expert System Analyses for the Geebung Creek and Peppermint Catchments, as well as a Combined Confusion Matrix for Both Catchments.

	Catchment		
Origin of Map	Geebung Creek	Peppermint	Combined
Conventional	78.3	70	73.6
Expert System	73.9	66.7	69.8

Table 8. The Overall Mapping Accuracy, within 95 Percent Confidence Intervals, of the Combined Confusion Matrix for Both Catchments.

Origin of Map	Combined Catchment
Conventional	60.6
Expert System	56.3

map form which we are conditioned to perceive as accurate. However, error is introduced by the interpreter, such as generalizing the polygon boundaries or transferring the analog line work on the aerial photographs to a digital vector format in the GIS.

The 53 independent sample points show that the overall accuracy of the expert system is only slightly less than that of the conventional map, i.e., 69.8 percent as opposed to 73.6 percent. Because the conventional map is considered operational by the Forestry Commission of New South Wales, and there is no statistical difference between the conventional map and the expert system, we conclude that the accuracy of the soil map produced by the expert system is also operational.

Using the technique proposed by Thomas and Allcock (1984), there is 95 percent confidence that 60.6 percent of the cells are correctly classified by the conventionally derived map, and 95 percent confidence that the accuracy of the expert system map is 56.3 percent. If a greater number of samples were obtained, the accuracy figure would of course rise. However, in order to establish confidence limits for each class, approximately 250 sample points would be required. This is a huge sample, given the constraints of access, and the time required to describe soil pits, in a remote natural forest.

An objection to the expert system approach (and also to conventional soil mapping!) is that the result is categorical, ignoring the gradual variation and measurement error in environmental data, especially with respect to soil. The creation of categorical input data layers from essentially continuous surfaces creates problems when generating rules. For example, consider the aspect layer. The "northerly" class could be delineated from the "north-easterly" class when the aspect becomes greater than 45°. An aspect of 5° is certainly more northerly than an aspect of 43°! Therefore, placing the input data into categories, coupled with the rather vague definitions of the soil-landscape units, makes mapping of forest soils by any method difficult.

Error may be introduced into a GIS by the input data being incorrect, or from errors inherent in the analysis procedure(s) used to model the data (Burrough, 1986). With an expert system, erroneous data may be introduced by the input data layers, or through the rules in the knowledge base. For example, the terrain models (slope and gradient) generated from a digital elevation model contain error, both as a result of incorrect *x,y,z* coordinates in the original digital elevation model, as well as a result of the assumptions of the algorithms used to calculate these terrain parameters (Skidmore, 1990). Heuvelink *et al.* (1989) developed techniques to propagate error in a GIS analysis by incorporating the error introduced by the GIS data layers, as well as the error from the regression models used in the analysis.

A number of examples are apparent from this study where incorrect input data layers caused the wrong soil landscape unit to be predicted by the expert system. The heath vegetation type is clearly associated with a particular soil type (AS2) (compare Plate 2 with Plates 5 and 6). There is a large area of AS2 towards the north of the conventionally derived soil map which does not appear on the expert system map (Plate 6). The explanation is that the heath type was not distinguished on the aerial photographs (or on the ground), so the DS2 soil type was predicted by the expert system. The soil scientist actually walked over the area, and discovered the AS2 soil landscape unit, but could not explain the occurrence of the soil in this location.

Thus, it appears that the expert system does not yield a totally accurate map. But it is interesting that the expert system and the conventional maps have the same statistical accuracy, indicating that the errors inherent for either method must be of the same magnitude.

Conclusion

In this study, we have shown that there is statistically no difference between the maps produced by an expert system linked to a GIS, and by an experienced soil scientist. The accuracy of the results generated by both methods is considered to be acceptable for forest managers. An interface between the GIS and the expert system has been developed and is presented. When the acceptable mapping accuracies are considered in conjunction with the mapping accuracy results, it is concluded that the expert system-GIS technology is operational. The successful integration of the expert system with a GIS, and the development of an interface, should encourage the use of this technique for a range of applications.

Acknowledgments

This work has been partly assisted through the Australian Research Council. The assistance of Dr. E.M. O'Loughlin and staff of the Division of Water Resources, CSIRO, in developing a TOPOG model over the study area is appreciated. Comments from anonymous referees assisted in preparing the manuscript.

Appendix

An Example of a Knowledge Base

```
!
! Knowledge base for forest soil mapping
! Note that an exclamation mark signifies a comment

fact(className,[RC, DS1, DS2, AS1, AS2]).

fact (priori_prob,[,[15, 15, 20, 20, 20]).

!
! Soil wetness-topographical position
! ==================================

! Abbreviations:
!
! topographical position:
! G:Gully
! L:Lower midslope
! M:Midslope
! U:Upper midslope
! R:Ridge
!
! wetness:
! V:Very Dry
! D:Dry
! M:Moist
! W:Wet
!
! e.g. GVD = gully very dry
! ------------------------------------------------------------------------
! probabilities of soil wetness index
! given classes.
!

fact(probE1,[40, 60, 20, 20, 30]). ! GVD
fact(probE2,[30, 50, 50, 40, 50]). ! GD
fact(probE3,[20, 40, 60, 50, 70]). ! GM
fact(probE4,[10, 40, 50, 40, 80]). ! GW
fact(probE5,[30, 60, 50, 40, 20]). ! LVD
fact(probE6,[30, 50, 60, 60, 30]). ! LD
fact(probE7,[20, 40, 70, 60, 40]). ! LM
fact(probE8,[10, 30, 60, 50, 60]). ! LW
fact(probE9,[40, 70, 50, 40, 20]). ! MVD
fact(probE10,[30, 60, 60, 60, 30]). ! MD
fact(probE11,[20, 50, 70, 60, 40]). ! MM
fact(probE12,[10, 30, 60, 40, 60]). ! MW
fact(probE13,[50, 70, 40, 30, 20]). ! UVD
fact(probE14,[50, 70, 40, 30, 20]). ! UD
fact(probE15,[40, 50, 70, 50, 40]). ! UM
fact(probE16,[30, 40, 60, 30, 60]). ! UW
fact(probE17,[60, 50, 20, 10, 10]). ! RVD
fact(probE18,[60, 50, 30, 10, 10]). ! RD
fact(probE19,[50, 50, 50, 20, 30]). ! RM
fact(probE20,[40, 40, 60, 30, 40]). ! RW
!
! probabilities of gradient given classes.
!
fact(probE21,[70, 20, 35, 65, 70]). ! <5
fact(probE22,[30, 60, 50, 40, 40]). ! 6-10
fact(probE23,[40, 60, 30, 30, 30]). ! 11-20
fact(probE24,[30, 60, 50, 30, 20]). ! >21
!
! probabilities of vegetation type given classes.
!
fact(probE25,[40, 60, 40, 30, 20]). ! STA/BLS
              ! Silvertop Ash/Blue Leaved Stringybark
fact(probE26,[30, 40, 50, 40, 20]). ! YS
              ! Yellow Stringybark
fact(probE27,[20, 40, 50, 55, 30]). ! Y/STA/S
              ! Yertchuk/STA/Stringybark
fact(probE28,[40, 50, 40, 30, 20]). ! STA/S/YS
              ! STA/Stringybark/Yellow Stringybark
fact(probE29,[10, 10, 20, 20, 20]). ! H
              ! Heath
fact(probE30,[30, 40, 50, 30, 20]). ! YS/MG
              ! Yellow Stringybark/Monkey Gum


fact(infile 1,"g_evibs.dat").
fact(out1,"gasb6.svf").
fact(out2,"gasb6.svf.label").

! RULES
! ==========
```

```
!

rule_1:if true
       then read nprData from infile1.

rule_2:if true
       then assign layer = 0 and
         using priori_prob to update_prob and
         assign evidence = 1 and
         assign layer = 1.

! ---------------------------------------------------------------------------
layer_1::
! ---------------------------------------------------------------------------
! soil wetness-topographical rules
!
rule_30: if layer > 1 or
         layer < 1
       then exit layer_1.

rule_31: if evidence = 1
       then using probE1 to update_prob and
         assign evidence = evidence + 1.

rule_32: if evidence = 2
       then using probE2 to update_prob and
         assign evidence = evidence + 1.

rule_33: if evidence = 3
       then using probE3 to update_prob and
         assign evidence = evidence + 1.

rule_34: if evidence = 4
       then using probE4 to update_prob and
         assign evidence = evidence + 1.

rule_35: if evidence = 5
       then using probE5 to update_prob and
         assign evidence = evidence + 1.

rule_36: if evidence = 6
       then using probE6 to update_prob and
         assign evidence = evidence + 1.

rule_37: if evidence = 7
       then using probE7 to update_prob and
         assign evidence = evidence + 1.

rule_38: if evidence = 8
       then using probE8 to update_prob and
         assign evidence = evidence + 1.

rule_39: if evidence = 9
       then using probE9 to update_prob and
         assign evidence = evidence + 1.

rule_40: if evidence = 10
       then using probE10 to update_prob and
         assign evidence = evidence + 1.

rule_41: if evidence = 11
       then using probE11 to update_prob and
         assign evidence = evidence + 1.

rule_42: if evidence = 12
       then using probE12 to update_prob and
         assign evidence = evidence + 1.

rule_43: if evidence = 13
       then using probE13 to update_prob and
         assign evidence = evidence + 1.

rule_44: if evidence = 14
       then using probE14 to update_prob and
         assign evidence = evidence + 1.

rule_45: if evidence = 15
       then using probE15 to update_prob and
         assign evidence = evidence + 1.

rule_46: if evidence = 16
       then using probE16 to update_prob and
         assign evidence = evidence + 1.

rule_47: if evidence = 17
       then using probE17 to update_prob and
         assign evidence = evidence + 1.

rule_48: if evidence = 18
       then using probE18 to update_prob and
         assign evidence = evidence + 1.

rule_49: if evidence = 19
       then using probE19 to update_prob and
         assign evidence = evidence + 1.

rule_50: if evidence = 20
       then using probE20 to update_prob and
         assign evidence = 1 and
         assign layer = layer + 1.

! ---------------------------------------------------------------------------
layer_2::
! ---------------------------------------------------------------------------
! slope gradient rules
!
rule_60: if layer > 2 or
         layer < 2
       then exit layer_2.

rule_61: if evidence = 1
       then using probE21 to update_prob and
         assign evidence = evidence + 1.

rule_62: if evidence = 2
       then using probE22 to update_prob and
         assign evidence = evidence + 1.

rule_63: if evidence = 3
       then using probE23 to update_prob and
         assign evidence = evidence + 1.

rule_64: if evidence = 4
       then using probE24 to update_prob and
         assign evidence = 1 and
         assign layer = layer + 1.

! ---------------------------------------------------------------------------
layer_3::
! ---------------------------------------------------------------------------
! forest overstory rules
!
rule_70: if layer > 3 or
         layer < 3
       then exit layer_3.
rule_71: if evidence = 1
       then using probE25 to update_prob and
         assign evidence = evidence + 1.
```

```
rule_72: if evidence = 2
      then using probE26 to update_prob and
        assign evidence = evidence + 1.

rule_73: if evidence = 3
      then using probE27 to update_prob and
        assign evidence = evidence + 1.

rule_74: if evidence = 4
      then using probE28 to update_prob and
        assign evidence = evidence + 1.

rule_75: if evidence = 5
      then using probE29 to update_prob and
        assign evidence = evidence + 1.

rule_76: if evidence = 6
      then using probE30 to update_prob and
        assign evidence = evidence + 1 and
        assign layer = layer + 1.

! ----------------------------------------------------------------------------
update_prob::
! ----------------------------------------------------------------------------
!

rule_90: if layer > 3
      then do update_prob and
        writ expClass to out1 and
        write className to out2 and
        exit update_prob.
```

References

Baur, G.N., 1965. *Forest Types in New South Wales,* Research Note No. 17, Forestry Commission of N.S.W., Sydney, Australia.

Beams, S.D., 1980. *Magnetic Evolution of the South-West Lachlan Fold Belt, South Eastern Australia,* unpublished PhD thesis, La Trobe University, Melbourne, Victoria, Australia.

Berry, B.J.L., and A.M. Baker, 1968. Geographic sampling, *Spatial Analysis: A Reader in Statistical Geography* (B.J.L. Berry and D.F. Marble, editors), Prentice-Hall, New Jersey.

Bishop, Y.M., S.E. Fienburg, and P.W. Holland, 1975. *Discrete Multivariate Analysis: Theory and Practice,* MIT Press, Cambridge, Massachusetts.

Burrough, P.A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment,* Clarendon Press, Oxford.

Cohen, J., 1960. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement,* 20(1):37–46.

Congalton, R.G., 1988. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data, *Photogrammetric Engineering & Remote Sensing,* 54(5):593–600.

Congalton, R.G., R.G. Oderwald, and R.A. Mead, 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques, *Photogrammetric Engineering & Remote Sensing,* 49(12):1671–1678.

Davis, J.R., J.R.L. Hoare, and P.M. Nanninga, 1986. Developing a fire management expert system for Kakadu National Park, Australia, *Journal of Environmental Management,* 22:215–227.

Davis, J.R., and G. McDonald, 1993. *Applying a Rule-Based Decision Support System to Local Government Planning,* Springer–Verlag, Berlin.

Forsyth, R., 1984. *Expert Systems: Principles and Case Studies,* Chapman and Hall, London.

Fung, T., and E. LeDrew, 1988. The determination of optimal threshold levels for change detection using various accuracy indices, *Photogrammetric Engineering & Remote Sensing,* 54(10):1449–1454.

Goodenough, D.G., M. Goldberg, G. Plunkett, and J. Zelek, 1987. An expert system for remote sensing, *IEEE Transactions on Geoscience and Remote Sensing,* GE-25(3):349–359.

Halpern, M., 1987. Turing's Test and the ideology of artificial intelligence, *Artificial Intelligence Review,* 1:79–93.

Hay, A.M., 1979. Sampling design to test land use accuracy, *Photogrammetric Engineering & Remote Sensing,* 45(4):529–533.

Heuvelink, G.B.M., P.A. Burrough, and A. Stein, 1989. Propagation of errors in spatial modelling with GIS, *International Journal of Geographical Information Systems,* 3(4):303–322.

Hutchinson, M.F., 1989. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits, *Journal of Hydrology,* 106:211–232.

Mackay, D.S., V.B. Robinson, and L.E. Band, 1992. Development of an Integrated Knowledge Based System for Managing Spatiotemporal Ecological Simulations, *GIS/LIS '92 Proceedings,* ACSM-ASPRS-URISA-AM/FM, San Jose, 2:494–503.

O'Loughlin, E.M., 1986. Prediction of surface saturation zones in natural catchments by topographic analysis, *Water Resources Research,* 22(5):794–804.

Parker, H.D., 1991. Commentary: The role of the private sector in GIS, *Photogrammetric Engineering & Remote Sensing,* 57(11): 1465–1466.

Rauscher, H.M., and T.M. Cooney, 1986. Using expert system technology in a forestry application: The CHAMPS experience, *Journal of Forestry,* 84(3):14–17.

Rosenfield, G.H., and K. Fitzpatrick-Lins, 1986. A coefficient of agreement as a measure of thematic classification accuracy, *Photogrammetric Engineering & Remote Sensing,* 52(2):223–227.

Ryan, P.J., 1993. Soil Formation on the Wallagaraugh Adamellite, south eastern N.S.W., Australia, *Catena,* 20:543–561.

Schank, R.C., 1988. What is AI, anyway? *AI Applications,* 2(2–3):12: 16.

Skidmore, A.K., 1989a. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model, *International Journal of Geographical Information Systems,* 3(4): 323–334.

———, 1989b. An expert system classifies eucalypt forest types using Landsat Thematic Mapper data and a digital terrain model, *Photogrammetric Engineering & Remote Sensing,* 55(10):1449–1464.

———, 1989c. Forest resources can be assessed using geographic information systems, remotely sensed data, and an expert system, *Forest Planning for People: Proceedings of the Thirtieth Biennial Conference of the Institute of Foresters of Australia,* Institute of Foresters of Australia, Leura, New South Wales, pp. 163–172.

———, 1990. Terrain position as mapped from a gridded digital elevation model, *International Journal of Geographical Information Systems,* 4(1):33–49.

Skidmore, A.K., P.J. Ryan, W. Dawes, D. Short, and E. O'Loughlin, 1991. Use of an expert system to map forest soils from a geographical information system, *International Journal of Geographical Information Systems,* 5(4):431–445.

Skidmore, A.K., J. Baang, and P. Luckananurug, 1992. Knowledge based methods in remote sensing and GIS, *Proceedings Sixth Australasian Remote Sensing Conference,* 2-6 November, Wellington, New Zealand, 2:394–403.

Stock, M., 1987. AI and Expert Systems: An Overview, *AI Applications,* 1(1):9:17.

Thomas, I.L., and G.M. Allcock, 1984. Determining the confidence interval for a classification, *Photogrammetric Engineering & Remote Sensing,* 50(10):1491–1496.

Tomlin, C.D., 1987. Introduction to Geographic Information Systems, *MAP Manual,* Yale School of Forestry and Environmental Studies, New Haven, Connecticut.

Weiss, S.M., and C.A. Kulikowski, 1984. *A Practical Guide to Designing Expert Systems,* Rowman and Allanheld, New Jersey.

(Received 6 December 1993; accepted 18 July 1994; revised 10 October 1994)