MANAGING KNOWLEDGE UNCERTAINTY IN AUTOMATED SYSTEMS FOR HANDLING SPATIAL DATA

A.J. Comber[1], A.N.R. Law[2], J.R. Lishman[3]

[1] Department of Geography, University of Leicester, Leicester LE1 7RH
email: ajc36@le.ac.uk

[2] Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen, AB15 8QH.

[2] Department of Computing Science, University of Aberdeen, Aberdeen AB24 3UE.

Abstract

Three commonly used techniques for combining uncertain evidence are explored with reference to different three types of land cover knowledge: numerical distributions, relative spectral distances, and human expert "rules of thumb". In attempting to combine such evidence Bayes', Dempster-Shafer and Endorsement theories answer different questions depending on nature of the land cover evidence (completeness and format). The approaches therefore have different utilities in the development of automated approaches to land cover monitoring. Whilst Bayes' and Dempster-Shafer theories may be more useful in situations where evidence is expressed numerically, Bayes' theorem requires a complete probability model. The advantage of Endorsement theory derives from its ability to represent different kinds of evidence in a natural form. It is a fundamentally symbolic approach that represents and reasons with knowledge of real-world problems and allows inferences to be drawn from partial knowledge. Such an approach is advantageous in knowledge acquisition and expert system development.

Keywords: uncertainty, land cover, evidence, Bayes Theorem, Dempster-Shafer, Endorsement Theory

1. Introduction

The aim of this work is to explore the relative utility of three methods for combing uncertain evidence in order to facilitate automated GIS approaches to land cover monitoring. Such approaches are being developed to update the Land Cover of Scotland 1988 survey (LCS88) through the integration of knowledge about land cover in an artificial intelligence (AI) environment (Comber et al, 2001). This knowledge includes land cover bio-geographic, and reflectance characteristics, expert descriptions of the land cover changes possible and interpreter "rules of thumb" used in the process of land cover mapping from aerial photography. Such interpretation routinely involves the use of contextual information (Paine, 1981). This may be from the imagery (the tone, texture and contrast within the scene), the bio-physical conditions at the micro- (hillside) or macro- (landscape) scale, evidence of the local land management, and from first hand knowledge of the area. The interpreter brings the combined effects of these factors upon the land cover together.

There is considerable uncertainty associated with such land cover knowledge: land cover classes cannot be uniquely defined in terms of their species composition (Comber et al, 2001), their position in bio-physical feature space (Armstrong and Milne, 1995), their reflectance characteristics (e.g. Price, 1994; Blackburn and Steele, 1999) and the rules of thumb employed to map them. In order to facilitate the development of expert land cover monitoring systems that incorporate data and knowledge from a variety of sources, such as SYMOLAC (Skelsey, 1997), a further dimension of uncertainty has to be considered: that of the rules used to combine the data. The knowledge used by human interpreters may be better represented as a series of "if X then *probably* Y" rules, with the *probably* having various degrees of strength depending on the land cover under consideration, the geographic location of the scene, the management context, and so on. Some of the land cover knowledge is described in qualitative terms and does not lend itself to easy and transparent numeric representation. It is difficult to ascribe numerical values to the strength of the evidence provided by the expert rules of thumb – their strength will depend on the specific context of the land cover change under consideration – and allocating values to qualitative statements would be open to criticism that it is not empirical. Any results derived by analysis of such evidence could be accused of being pre-determined. It is within this application complexity that methods are sought to manage the uncertainty from different knowledge sources.

Three methods for combining uncertain information commonly are described in the AI literature: Bayes', Dempster-Shafer and Endorsement theories. Of these, Bayes' and Dempster-Shafer are the most widely known and have caused two distinct schools of thought to be delineated. Bayes' theorem is described in most introductory texts to AI (e.g. Jackson, 1986; Norvig and Russell, 1995), and see Howson and Urbach (1993) for a more detailed discussion. Much recent work describes modifications to Dempster-Shafer theory, but Parsons (1994) provides a clear introduction to the application and mechanics of Dempster-Shafer. A good description of the arguments and counter-arguments put forward by both sides of the Bayes-Dempster-Shafer dichotomy is contained in a text edited by the main protagonists from either side: Shafer & Pearl (1990). Endorsement theory is the third approach to combining uncertain information described here. It is non-numeric approach was developed by Cohen (1985), and has been used in some automated mapping applications, where different types of evidence have been combined (e.g. Srinivasan and Richards, 1990; Srinivasan and Richards, 1993; Skelsey, 1997).

Section 2 describes the land cover knowledge, the three approaches to combining uncertain evidence are introduced in Section 3 and Section 4 discusses the issues raised by apply Dempster-Shafer, Bayes and Endorsement theorems to the different types of land cover knowledge. Some conclusions are presented in Section 5.

2. Knowledge or "evidence"

Different analyses have elicited various types of knowledge and data about different aspects of land cover. These include matrices of the possible land cover changes at *Time2* from any land cover class at *Time1*, distributions of land cover over different environmental gradients, land cover class reflectance characteristics and statements from air-photo interpreters of their rules of thumb. It is desirous to combine this information in order to determine the likely land cover class of an area identified as having changed since *Time1*. However this evidence has "uncertainty" associated with it.

The transition matrices provide a list of all the *possible* land cover changes at *Time2* from any land cover class at *Time1*. This list forms the initial set of land cover change direction hypotheses and is comprehensive. However it includes some extremely unlikely transitions such as those that will not occur because of the time frame for the transition. For instance, changes to Peatland vegetation, whilst possible, will take millennia.

The rules of thumb can be specific or ambiguous in their applicability. This ambiguity can have an ecological basis and can be rooted in the nature of the classification scheme, LCS88. They vary in the confidence with which they can be applied. For instance, rules such as "There will not be any changes in scattered-rock status (an LCS88 classification feature)" could be reliably used to eliminate certain change hypotheses either pre- or post-analysis. Whilst one would have much less confidence in applying a rule such as "There *probably* will not be any changes in scattered-tree status (another LCS88 feature)". Other rules of thumb define land cover bio-geographic parameters such as "Undifferentiated Smooth Grasslands will only be found on slopes greater than 15°".

The presence of any particular vegetation community can be seen as the interaction between the environmental gradients and the management practice at that point (Comber et al 2001). Therefore bio-geographic phenomena might be expected to indicate the presence of land cover classes. Overlaying environmental datasets provides distributions of land cover over such gradients. An example of the distribution of land cover classes with slope is shown in Table 1. The land cover classes in Table 1 are a subset of the actual classes in the Elgin-Speyside area in NE Scotland. Whilst the distribution of land cover with slope is presented here, it is noteworthy that land cover class is not uniquely defined by its bio-geographic position. This contributes uncertainty about the extent to which such environmental data can be used to individual land cover classes.

Further land cover knowledge is provided by comparing the reflectance characteristics of the change area with those of the various land cover classes. The "closeness" of the change area reflectance properties to the class populations can be measured and this provides another source of evidence about land cover change directions. This can be done by measuring the relative distance of the change area median to some measure of central tendency such as the class medians. In Table 2, this relative distance is derived from the inverse of the measured distance. This calculation of spectral distance is very simplistic. However it serves to illustrate the type of data and information about the spectral characteristics of land cover classes might be available as evidence, regardless of remote sensor, band or band index. The data in Table 2 illustrates the origins of this work: land cover class reflectance values are not unique and it would difficult to label the change area to any particular land cover class with confidence (although it could be stated that it likely to be a grassland).

In summary, the descriptions of the data and knowledge have illustrated the uncertainty involved in using such evidence to infer land cover change direction. Indeed, some of the evidence may be contradictory. For instance, from the data in Table 1 it is apparent that some Undifferentiated Smooth Grasslands is found on slopes less than 15°, contrary to one of the expert statements. Further, not all of the *possible* land cover change directions indicated by the transition matrix are likely. There may also be doubt about the applicability of interpreter rules of thumb, in the strength of evidence provided by distributions land cover across bio-geographic phenomena and by reflectance characteristics for determining the direction of land cover change. However, together these pieces of evidence, although uncertain, can provide enough weight to indicate *actual* land cover change direction. What is needed is some method to combine such uncertain evidence. It should be emphasised that this uncertainty is not associated with the cartographic paradigm, commonly considered in much GIS research. We are not concerned here with such issues as boundary position and polygon attribute allocation. Rather, it is the uncertainty in the land data and knowledge itself and the extent to which it can be reliably applied to the problem of determining land cover change direction that is explored via three techniques for combining uncertainty evidence.

3. Approaches to combining uncertain evidence

3.1. Bayes' Theorem

Bayes' theorem provides a simple numerical method for updating the probability of a hypothesis given an observation of evidence. It computes the probability of an hypothesis or event, *h* given the evidence, *e* in support of that event, *P(h|e)*. More complex forms exist for revising the probability of *a number* of hypotheses by pooling the weights of pieces of evidence (e.g. Cohen, 1985). However the issues raised by using this approach to combining the evidence described in Section 2 can be illustrated using the simpler, more familiar form:

$$P(h \mid e) = \frac{P(h) * P(e \mid h)}{P(e)}$$

Equation 1.

where, *p(h|e)* is the posterior probability of hypothesis *h* given evidence *e*,

p(h)) is the prior probability,

p(e|h) is the likelihood, and

p(e) is the probability of the evidence.

As we have no initial reason to believe one hypothesis over another *p(h)*, nor any belief in the probability of the evidence, *p(e)*, these terms can considered constant for all hypotheses. Therefore the relative value posterior probability for each hypothesis depends on the likelihood of the evidence given the hypothesis, *p(e|h)*. From the areas of different land cover classes with different categories of slope data in Table 1 we can calculate *p(e|h)* on a *per slope category* basis, that is by dividing each element in the table by total area for that type of evidence. These are shown in Table 3.

The issues raised by using Bayes' theorem to determine the strength of belief in each hypothesis can be illustrated by considering the data in Table 3. If the change area has a slope in the range 3-8°, then the evidence in Table 3 would indicate a shift to Arable (0.45) despite only a third or so of the total Arable area being in that category of slope. This is due to the large area of Arable in the Elgin-Speyside region. Other classes such as Undifferentiated Heather Moorland, have a low probability (0.11) despite a third of its total area being on that type of slope. This is after all what Bayes' theorem does: it determines the probability of the hypothesis given the evidence. The entities used to provide evidence in this case are the areas of land cover class in each slope category. The hypotheses with large areas will be classes with the highest probability and indicated as the likely land cover change. (If counts of land cover polygons in each slope category were the entities, then the results of applying Bayes' theorem to such data would be biased towards land cover classes with largest number of polygons.) However, this is not necessarily the answer that we require. What we would like is the probability of the change class being Undifferentiated Heather Moorland given the evidence that slope is 3-8°, relative to the distribution of Undifferentiated Heather Moorland across the categories of slope. This would mean calculating the distributions of land cover classes with environmental features on a *per land cover class* basis. However, to organise the data in this way, so that any results are not biased towards the land cover classes with the largest entities (in this case area), changes the question that is being answered. Organising the data so that it reflects the relative likelihood of the evidence given the hypothesised land cover class, changes the question that is answered. By calculating distributions on a per land cover class basis means that question being answered is "*how likely is the evidence (over other types of evidence) given the class?*" rather than "how likely is the class (over other classes) given this evidence?".

The issues raised by the organisation of the slope data are compounded by the application of Bayes' theorem to the other evidence. The reflectance "distances" in Table 2 do not represent a distribution across all the classes: they are not probability measures summing to unity. Rather they are measures of the closeness of the change area to the possible change directions. In that sense they are *per class* evidence and indicate the likelihood of the evidence given the hypothesis (land cover class). If the interpreter rules of thumb were to be incorporated into a Bayesian analysis they would have to be given numeric values. For those with strong conclusions, a value can be allocated to the *x* number of hypotheses that are supported by the rule of 1 / *x*. The allocation of values to the more ambiguous rules of thumb is more problematic. Consider the rule that "There *probably* will not be any changes in scattered-tree status". We have no indication of the strength of the rule relative to any of the change hypotheses. Therefore allocating a value would be arbitrary.

3.2. Dempster-Shafer

Dempster-Shafer theory is another numerical method for combining uncertain evidence. A numerical measure of the weight of evidence (mass assignment, *m*) is assigned to *sets of hypotheses* as well as individual hypotheses. Dempster-Shafer is not a method that considers the evidence hypothesis by hypothesis as Bayes' theorem does, rather the evidence is considered in light of the hypotheses. A second piece of evidence is introduced by combining the mass assignments (*m* and *m'*) using Dempster's rule of combination, to create a new mass assignment *m''*. Dempster's rule of combination is defined by:

$$m''(C) = \sum_{\substack{i,j \\ A_i \cap B_j = C}} m(A_i)m'(B_j)$$

(Equation 2, from Parsons, 1994)

That is, the combined mass assignment of C ($m''(C)$) is equal to the sum of $m(A_i) * m'(B_j)$ for all i and j such that set $A_i \cap B_j$ is equal C. The result of combining two assignments is that for any intersecting sets A and B, where A has mass M from assignment m and B has mass M' from assignment m', the belief accruing at their intersection is the product of M and M'. Or, "sum for each combination of A and B that overlap with C". This process can be repeated for additional evidence D, such that $m''''(E)$ equals each overlapping combination of $m''(C)$ and $m'''(D)$.

Applying Dempster-Shafer to the data or evidence we have presented is not straightforward. The data has to be organised so that a piece of evidence is seen to support a set of change hypotheses, with a percentage certainty that the solution is to be found in that set. For instance, if the change area has a slope in the range 3-8°, from Table 3 it would be difficult to determine the members of the subset of hypotheses supported by that evidence and the weight of evidence, *m*, for that subset. We would like to be able to say, for instance, that the subset of {Wet Heather Moorland, Undifferentiated Heather Moorland, Undifferentiated Smooth Grassland} is indicated by that evidence with a strength of *X*, as these classes have the largest proportion of their areas on slopes of that type (from Table 1). But because of the detail contained in Table 1, we also know that the subset members have considerable portions of their populations on other slope types. It would be preferable to have evidence with a greater degree of disaggregation such as a statement from an expert that they are 60% sure that evidence indicates one of the subset of hypotheses. In this sense there may be too much evidence in the data contained in Tables 1 and 3 for Dempster-Shafer. Of course, the data in Table 1 can be re-organised so that for each land cover class only one type of slope is indicated. This is shown in Table 4. Determining the subset of hypotheses from Table 4 is straightforward. However, further contortions are necessary to generate a mass assignment for evidence. One solution is to determine the mass assignment by the proportion of hypotheses supported by the (disaggregated) data. So that for slope type 3-8°, the subset {Wet Heather Moorland, Undifferentiated Heather Moorland, Undifferentiated Smooth Grassland} is supported with a strength of 3/7.

To derive similar subsets and strengths of belief for the reflectance data in Table 2, would require applying a threshold. We might decide that land cover classes whose medians are within 5 of the change area will constitute the subset of hypotheses supported by this evidence. The mass assignment for the subset could be calculated in the same manner as for the slope evidence, by the proportion of the total number of hypotheses falling within that threshold. The application of uncertain rules of thumb in a Dempster-Shafer approach requires that they be given numeric values. Again, similar to Bayes' theorem, we have no indication of the strength of the rule relative to any of the change hypotheses and any allocation of a numerical value to that rule would be arbitrary.

3.3. Endorsement Theory

Endorsement theory (Cohen, 1985) is a non-numerical approach to combining uncertain evidence. Evidence is given a description of the belief that it contributes to a hypothesis being true. Different strengths of beliefs are defined, and a hypothesis is given an endorsement depending on all the different statements of belief and disbelief, accrued from the evidence. Although endorsements and beliefs are given names, the names are tokens. Their meanings derive from specifying the situations in which they are applicable, how they combine and how they are ranked. This requires the definition of four aspects of the problem (Sullivan and Cohen, 1985):

i) Beliefs must be identified and named. For instance we could define and rank our beliefs and disbeliefs as:
   - "*conclusive-belief*" where a single piece of evidence alone indicates that the hypothesis is true;
   - "*prima-facie*" where the evidence alone would support the hypothesis, but may be contradicted;
   - "*strong*" where the evidence contributes some of the overall support for the hypothesis;
   - "*weak*" where the evidence supports the hypothesis, but not enough to conclude about it.

ii) The interaction of beliefs when combined must be specified to produce overall endorsements. In this example the endorsements are defined as:
   - "*definite-hypothesis*" when the evidence provides conclusive belief and no conclusive disbelief;
   - "*confident-hypothesis*" when the combined evidence provides prima-facie belief and no prima-facie disbelief;
   - "*likely-hypothesis*" when strong belief is greater than strong disbelief;
   - "*indicated-hypothesis*" when weak belief is greater than weak disbelief;
   - "*contradicte- hypothesis*" when the weights of belief and disbelief are equal, and further evidence is required to be able to conclude about the hypothesis.

iii) A system for ranking endorsements must be specified. The endorsements strength, in this example, goes from definite to confident to likely to indicated.

iv) Rules must be defined to decide when evidence accrued is "believable enough". This decision depends on the endorsement of the evidence and on the intended use of that evidence. It may be that a single piece of evidence providing conclusive belief in a hypothesis, would be enough to stop the reasoning process, and enable conclusions to be drawn. It is noteworthy that this kind of evidence is rare for the determination of land cover change direction.

From the above it is obvious that individual problem domains will have characteristic kinds of beliefs endorsements, and criteria for reasoning with them. Endorsement model makes sources of uncertainty explicit and takes a much more heuristic approach to reasoning about uncertainty than either Dempster-Shafer or Bayesian approaches. Firstly it is able to represent common knowledge (such as ambiguous API knowledge) in a natural form. Secondly its symbolic approach is adequate to represent and to reason with such knowledge and real-world problems. Thirdly such reasoning allows inferences to be drawn from partial knowledge (Srinivasan and Richards, 1993). The question that the endorsement-based approach seeks to answer is "what are the sources of uncertainty in the reasoning process, and where were they introduced?" The meaning of this answer is then interpreted through the method by which endorsements combine and how they are ranked.

These aspects of the approach can be illustrated by considering the slope evidence. Suppose that we again have a change area with a slope in the range 3-8°. A simple approach would be to take the evidence from Table 1 in isolation from other available evidence and to allocate beliefs to that evidence. One way of doing this is to go back to the raw data in Table 1. If the largest land cover area is also in the slope category 3-8°, then beliefs can be allocated to the hypotheses that change is to Wet Heather Moorland, Undifferentiated Heather Moorland and to Undifferentiated Smooth Grassland. We have no particular indication of the strength of evidence that slope type contributes to individual hypotheses, so one approach is to relate the strength of belief to the proportion of it in the same slope category as the change area. For instance, conclusive belief if the percentage of that class found on that slope type was greater than 95%, prima-facie belief if greater than 60%, strong belief if greater than 40% and weak belief if greater than 25%. The *per land cover class* distributions of slope type, and the associated strength of belief that they confer upon the evidence, according to the scheme above, are presented in Table 5.

The evidence from the reflectance data in Table 2 can also be allocated beliefs according to a scheme based on the proximity of the change area median to the class medians. For the interpreter rules of thumb we do not have the problem of trying to quantify them numerically. Due to the heterogeneous nature of much semi-natural land cover in terms of its bio-geographic distribution and definition, few rules of thumb will offer very strong or *prima-facie* evidence, and even fewer will contribute evidence that is *conclusive*. Those that do are self-defining. For instance, the rule of thumb that states "There will not be any changes in scattered-rock status". Most will contribute some *strong* evidence, which together with other contextual evidence enables the land cover to be determined. For instance, "Undifferentiated Smooth Grasslands will only be found on slopes greater than 15°". In this way the interpreter rules of thumb can be readily combined using Endorsement theory, which combines those qualitative terms explicitly.

4. Discussion and conclusions

The question that the Bayesian approach is answering is "what is the belief in A?" as expressed by the unconditional probability that A is true given evidence, e ?" In principal the Bayesian approach can be applied to any problem involving uncertainty, assuming that precise probabilities can be assessed for all events. The rules of the probability calculus are uncontroversial, they ensure that the conclusions are constant with these assessments and it is easy to understand what Bayesian theory does. It is the assumption of precision that can be unacceptable because in practice it can be difficult to make many precise assessments of probabilities. In order to apply Bayes' theorem in a consistent manner to the different types of evidence presented here requires some data re-organisation. The slope distributions have to be calculated on a *per land cover class* basis to avoid biasing results towards hypotheses with the largest number of entities (areas). The reflectance data has to be re-organised so that the distances summed to unity and the ambiguous rules of thumb have to be allocated numeric values. However, such contortions change the meaning of the data and the results. The conclusion is that Bayes' theorem is most suited to problems where there *are* probabilities for all events. It does poorly where there is partial or complete ignorance, or limited or conflicting information and it cannot deal with imprecise, qualitative or natural language judgements such as "if *A* then *probably B*". These are not easily accommodated because of the requirement for precise assessments and a complete probability model.

Whereas a Bayesian approach assesses probabilities directly for the answer, the Dempster-Shafer approach assesses evidence for related questions. It can be thought of as answering the question "what is the belief in A, as expressed by the probability that the proposition A is *provable* given the evidence?" (Pearl, 1988). In this sense Dempster-Shafer can model various types of partial ignorance, limited or conflicting evidence and is a more flexible model than Bayes' theorem. Although because of this, it has been criticised for not giving clear guidance to understand its conclusions (e.g. Walley, 1996; Howson and Urbach, 1993) and it can produce conclusions that are counter-intuitive. Dempster-Shafer is computationally simpler than Bayes' theorem. However some important types of uncertainty such as judgements of probability in ordinary language are not easily modeled. An approach for disaggregating the slope data into a single statement for each of the land cover classes was described, and a method found for allocating mass assignments. However, whilst it may be possible to re-organise the data and knowledge about land cover into a format suitable for combination, a considerable amount of evidence is lost in this process. In conclusion, Dempster-Shafer is most suited to situations where beliefs are numerically expressed and where there is some degree of ignorance, i.e. there is an incomplete model.

Endorsement theory allows the definition of beliefs and their interaction to be specified according to the problem domain being considered. The question that the endorsement-based approach answers is "what are the sources of uncertainty in the reasoning process, and where were they introduced?" The meaning of this answer is then interpreted through the method by which endorsements combine and how they are ranked. Despite some aspects of endorsement models being cumbersome they have one distinct advantage over numerical approaches: their results contain explicit information about *why* one believes and disbelieves. Consequently it is possible to reflect on these, and decide how to act - a very useful property for expert system development. Historically, expert systems have used numerical techniques to assess subjective degrees of belief in uncertain alternatives: it is very easy to rank them, and combinations are calculated by simple arithmetic rules. The application of endorsement-based approaches is most suited to situations where subjective degrees of belief do not generally behave as probabilities. The knowledge elicitation phase of the construction of expert systems is one such application area: domain experts are often uncomfortable committing themselves to numerical values. Numbers may be ambiguous and composed of salience and probability considerations. Endorsements are records of sources of uncertainty and provide explicit records of the introduction of uncertainty into the reasoning process. However, they may be inappropriate for domains in which numerical degrees of belief have a clear semantics and are adequate expressions of all information about uncertainty. In these situations Bayesian approaches may be preferable.

In conclusion, the relative value of Endorsement theory over Bayes' and Dempster-Shafer theories is due to:
1) Its ability to incorporate qualitative knowledge such as rules of thumb into the method of evidence evaluation;
2) Conflicting evidence is not aggregated and lost, as in numeric methods, rather it is explicitly handled and included in the hypothesis endorsement explanations;
3) It does not rely on mathematical functions to produce a numerical description of the strength of the combined evidence. Instead, Endorsement theory produces a description of the evidence, its strength of belief and the method by which it was combined, allowing the operator to trace back through the reasons for every hypothesis endorsement.

The utility of Endorsement theory for reasoning about uncertainty in AI applications therefore stems from its ability to represent different kinds of knowledge in a natural form. It is a fundamentally symbolic approach that represents and reasons with knowledge of real-world problems. It allows inferences to be drawn from partial knowledge and consequently the approach can function in the early stages of knowledge acquisition and expert system development.

ACKNOWLEDGEMENTS
This work was done as part of a Ph.D. thesis at the Macaulay Land Use Research Institute and the Computing Science Department, Aberdeen University. Thanks to MLURI and the Scottish Executive Rural Affairs Department for financial support through the Non-commissioned Research Programme.

References

Armstrong, H.M. & Milne, J.A., (1995), The Effects of Grazing on Vegetation and Species Composition. In D.B.A. Thompson, A.J. Hester &M.B. Usher (Eds), *Heath and Moorland Cultural Landscapes* (pp. 162-173). Edinburgh: HMSO.

Cohen, P.R., (1985). *Heuristic Reasoning About Uncertainty: An Artificial Intelligence Approach*. Boston: Pitman Advanced Publishing.

Comber, A.J., Law, A.N.R., & Lishman, J.R., (2001). Methodologies and Approaches for Automated Land Cover Change Detection. In P. Halls (Ed), *Spatial Information and the Environment (Innovations in Gis, 8)* (Chapter 2). London: Taylor and Francis.

Howson, C. & Urbach, P., (1993). *Scientific Reasoning: the Bayesian Approach*. Peru, Illinois: Open Court Publishing Company.

Jackson, P., (1986). *Introduction to Expert Systems*. Wokingham: Addison-Wesley.

Parsons, S., (1994). Some Qualitative Approaches to Applying the Dempster-Shafer Theory, *Information and Decision Technologies* 19(4), 321-337.

Pearl, J., (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo: Morgan Kaufmann.

Price, J.C., (1994). How Unique are Spectral Signatures? *Remote Sensing of Environment*, 49(3), 181-186.

Shafer, G. & Pearl, J., (1990), *Readings in Uncertain Reasoning*. San Mateo: Morgan Kaufmann.

Skelsey, C., (1997). *A System for Monitoring Land Cover*. Aberdeen University: Thesis, Ph.D. or available by prior arrangement at (http://bamboo.mluri.sari.ac.uk/SYMOLAC/). Contact Chris Skelsey at (chris_skelsey@yahoo.com).

Srinivasan, A. & Richards, J.A., (1990). Knowledge-Based techniques for Multi-Source Classification. *International Journal of Remote Sensing*, 11(3), 505-525.

Srinivasan, A. & Richards, J.A., (1993). Analysis of Spatial Data Using Knowledge-Based Methods. *International Journal of Geographic Information Systems*, 7(6), 479-500.

Sullivan, M. & Cohen, P.R., (1985). An Endorsement-Based Plan Recognition Program. In J.K. Aravind *Proceedings of the 9th International Joint Conference on Artificial Intelligence* (pp. 475-479). Los Angeles: Morgan Kaufmann.

Walley, P., (1996), Measures of uncertainty in expert systems. *Artificial Intelligence*, 83(1), 1-58.

Table 1. Land cover class distribution of slope in the Elgin-Speyside area (in 10,000m^2)

	Slope				
Land cover class	0-2°	3-8°	9-15°	16-25°	> 25°
arable	17209.38	9164.81	2218.43	51.80	1.25
dry heather moorland	919.70	2084.67	2456.46	445.97	11.91
wet heather moorland	520.21	548.54	131.65	0.00	0.00
undifferentiated heather moorland	1565.94	2195.54	1856.51	335.56	6.67
smooth grassland with rushes	2975.60	2975.36	216.85	13.52	0.05
smooth grassland with scrub	828.29	600.12	421.58	41.19	0.16
undifferentiated smooth grassland	378.20	475.53	453.84	91.47	4.34
Total	26007.61	20344.09	7791.39	982.22	24.56

Table 2. A measure of the distance between the class histogram medians and the histogram median of an area of change in Landsat band 2 for the Elgin-Speyside area

Land cover class	median	distance	1/distance
Change area	32		
arable	30	2	0.500
dry heather moorland	20	12	0.083
wet heather moorland	24	8	0.125
undifferentiated heather moorland	22	10	0.100
smooth grassland with rushes	30	2	0.500
smooth grassland with scrub	31	1	1
undifferentiated smooth grassland	34	2	0.500

Table 3. The likelihoods of the hypotheses given the slope evidence from Table 1.

	Slope				
Land cover class	0-2°	3-8°	9-15°	16-25°	> 25°
arable	0.66	0.45	0.28	0.05	0.05
dry heather moorland	0.04	0.10	0.32	0.45	0.48
wet heather moorland	0.02	0.03	0.02	0.00	0.00
undifferentiated heather moorland	0.06	0.11	0.24	0.34	0.27
smooth grassland with rushes	0.11	0.15	0.03	0.01	0.00
smooth grassland with scrub	0.03	0.03	0.05	0.04	0.01
undifferentiated smooth grassland	0.01	0.02	0.06	0.09	0.18

Table 4. The slope types containing the largest area for land cover classes in the Elgin-Speyside area are indicated by a "X".

	Slope				
Land cover class	0-2°	3-8°	9-15°	16-25°	> 25°
arable	X				
dry heather moorland			X		
wet heather moorland		X			
undifferentiated heather moorland		X			
smooth grassland with rushes	X				
smooth grassland with scrub	X				
undifferentiated smooth grassland		X			

Table 5. Per land cover class distributions of slope type, and the strength of belief that they confer.

	Slope				
Land cover class	0-2°	3-8°	9-15°	16-25°	> 25°
arable	0.60 *prima-facie*	0.32 *weak*	0.08 *none*	0.00 *none*	0.00 *none*
dry heather moorland	0.16 *none*	0.35 *weak*	0.42 *strong*	0.08 *none*	0.00 *none*
wet heather moorland	0.43 *strong*	0.46 *strong*	0.11 *none*	0.00 *none*	0.00 *none*
undifferentiated heather moorland	0.26 *weak*	0.37 *weak*	0.31 *weak*	0.06 *none*	0.00 *none*
smooth grassland with rushes	0.48 *strong*	0.48 *strong*	0.04 *none*	0.00 *none*	0.00 *none*
smooth grassland with scrub	0.44 *strong*	0.32 *weak*	0.22 *none*	0.02 *none*	0.00 *none*
undifferentiated smooth grassland	0.27 *weak*	0.34 *weak*	0.32 *weak*	0.07 *none*	0.00 *none*