UNCERTAINTY, VAGUENESS AND INDISCERNIBILITY: THE IMPACT OF SPATIAL SCALE IN RELATION TO THE LANDSCAPE ELEMENTS

A.J. Comber[1], P.F. Fisher[2], A.Brown[3],

[1.] Department of Geography, University of Leicester, Leicester, UK. E-mail: ajc36@le.ac.uk;
[2.] The giCentre, Department of Information Science, City University, London, UK e-mail: pff1@city.ac.uk;
[3.] Alan Brown, Countryside Council for Wales, Bangor, LL57 2LQ, UK. a.brown@ccw.gov.uk

KEY WORDS: uncertainty, scale, grain, habitats

ABSTRACT

Countryside agencies in the UK are interested in how to generate a range of Boolean maps whilst controlling uncertainties in ways that are appropriate to different landscape questions. This work illustrates the different effects of Bayes and Dempster-Shafer in combining ancillary information to augment remote sensing analyses. The example of translating between different habitat classifications is used (Phase I, Priority and Annex I) to test the hypothesis that it is possible to move beyond from the concept of crisp mappings to one based on bounded belief. The results highlight the importance of data scale and grain of process in the development context-sensitive Boolean maps and indicate that conservation managers need to be able to identify the 'best' decision in the face of uncertainty relating to the nature of the object or process and scale.

1. INTRODUCTION

Countryside agencies such as the Countryside Council for Wales (CCW) are responsible for reporting on and monitoring the rural environment. CCW are increasingly being asked to monitor the landscape pressures and effects relating to a series of drivers such as agri-environmental impacts, climate change and changes to structural support for farmers. Countryside agencies would like to be able to describe the landscape under a range of different policy initiatives. These include the traditional environmental roles relating to land cover habitats (e.g. Annex 1, Priority Habitats), but increasingly relate to such new questions. Each of these has their own set of constructs within which the landscape is viewed.

The problem addressed in this paper is how to translate different habitat classifications from existing ones, given some additional information (e.g. field survey, other data, remote sensing information). CCW has national Phase I habitat data (JNCC, 2003), but would like to be able to describe the landscape in terms of other habitats with different grains as a result of EU and national biodiversity legislation:

- Priority habitats as described in UK Biodiversity Action Plan (UK Government, 1994);
- Annex I habitats from the EC Habitats Directive which lists important high-quality conservation habitat types and species in its annexes (Commission of the European Communities, 1992);
- Phase II or National Vegetation Classification (NVC) habitats as specified by Rodwell (2006).

Traditionally, conservation agencies use their understanding of habitat semantics to integrate data: habitat A is translated into habitat B by considering the range of attributes or vegetation sub-classes classes within A and in B. However this process has occurred more or less covertly. Examination of data semantics allows sets of relations between classes to be constructed in light of the grain of process. The translation from Phase I to Annex 1 habitats for example, represents a refinement in grain. A scheme of this integration process is shown in Figure 1.

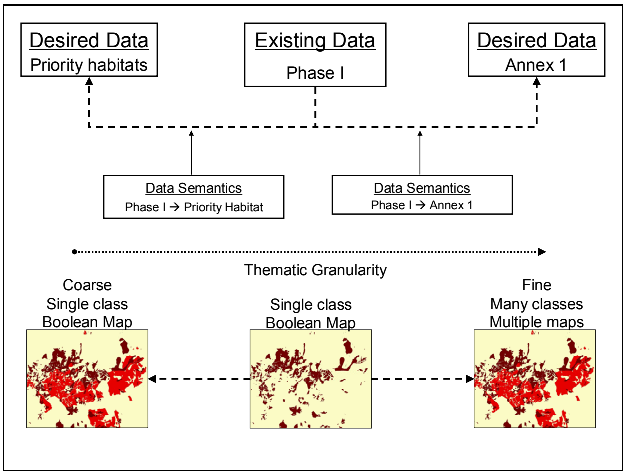

Figure 1. Issues in translating between different habitat classifications based on data semantics.

2. BACKGROUND

Countryside agencies are faced with two problems: First, how to translate information from their existing data holdings to answer new landscape questions. The data may be thematically or spatially coarser than would be ideal to answer these. Second, it is difficult to incorporate the uncertainties associated with data translations. This is essential as the any uncertainty involved will necessarily depend on the question being asked of the data (Comber et al., 2006, 2004a).

The multiplicity of questions that may be asked of any dataset raises a number of issues: 1) How to generate a range of possible maps which manipulate the data (e.g. fusions and aggregations) in different ways; 2) How to choose the most appropriate map for the task in hand (i.e. what to display)? 3) How to understand and quantify uncertainty that relates to this specific application?

A recent mapping initiative sought to evaluate how updated maps of Phase I habitats in Wales could be reworked to answer other questions at different scales and granularities. This paper considers the uncertainty of feature representation where a number of habitats could be identified at any particular point, and how beliefs and preferences can be incorporated in a consistent way into the final map. The resulting maps are called *context-sensitive*, because they have been produced to meet a particular need.

Bayes and Dempster-Shafer are applied to two example questions relating to different habitat granularities and scales for which practical management (specifically the monitoring of burning activity) requires decisions that relate to patch size and landscape context. For example there may be patches of bog within the upper bound of potential bog which are too small to be managed independently of the surrounding heath. Similarly the potential upper bounds may lie beyond those patches assigned to the habitat class so that patches of heath are treated as bog because their mosaic with bog is too intimate, in which case bog management takes priority over heath management.

Both examples can be considered in from a legal or a conservation perspective. The legal one relates to the legitimacy of burning activity and the conservation one relates to monitoring of important (Annex I) habitats. In legal situations Bayesian or probabilistic approaches are more appropriate where there is less uncertainty about the evidence and there is a need to identify the *probable* outcomes, as they may have implications (e.g. prosecutions). Where the evidence has more uncertainty, then approaches that identify upper and lower bounds are more appropriate as they explicltly incorporate the uncertainty by showing the *possible* extent of different habitats.

3. UNCERTAINTY

All land cover maps incorporate some uncertainty, even if this is not obvious: Error and uncertainty arise at every stage in the production of maps from remotely sensed imagery (Fisher, 1997; Comber et al., 2005a, 2005b). Remote sensing of land cover is predicated on the assumption that the land cover features of interest can be statistically separated and discerned from remotely sensed imagery. Most land cover datasets are Boolean classifications which allocate each data object (pixel, parcel) into one class and membership of any class is binary. There are uncertainties associated with process of mapping land cover from remotely sensed imagery, relating to:
- The discerning power of the image may not be able to identify land cover at the required level of grain (spatial resolution);
- The target land covers may themselves not be spectrally homogenous (spectral resolution);

The end result is that statistical clusters in *N* image bands may not relate to the desired or target land covers resulting in class to class confusions. These issues are well described in the literature: Freidl et al (2001) describe issues relating to spatial resolution; Comber et al. (2004b) spectral resolution, but are rarely accommodated operationally where the end result is that a Boolean allocation decision is made for each object and any uncertainty is often conveniently ignored.

There are a number of issues with this land cover mapping model:
1) Land covers may be composed of heterogeneous mixtures of vegetation which may be beyond the spectral and spatial resolution of the remotely sensed data. This is often the case in upland semi-natural landscapes;
2) Many land cover initiatives seek to augment analyses of remotely sensed data with other information;
3) Land cover maps are used for many other purposes than that for which they were originally constructed and are used to answer multiple landscape questions, not just the extent and distribution of habitats, such as Phase I.

Therefore, there is a growing interest in being able to re-allocate data objects into different classes for different landscape questions: context sensitive maps. The re-allocation may be based on the uncertainty associated with the original Boolean allocation and/or due to different weights being given to the supporting evidence, for instance from ancillary data.

Most approaches to managing uncertainty in the GIS and the nature conservation communities adopt a probabilistic approach under the assumption that the various pieces of data and evidence are independent (i.e. they are not correlated with other data or evidence). This is problematic for a number of reasons. First, the much environmental data is spatial auto-correlated. Second, the classic error assessment method, tabulating predicted against observed in a correspondence matrix, assumes that like is being compared with like. This is not the case. Field surveys relate to land cover to plant communities, whilst remotely sensed classes exist in spectral or image band feature space. These are fundamentally different mental constructs of land cover (see Comber et al 2005b for a full description). Third, the landscape objects themselves are assumed to well defined (i.e. not vague, indeterminate or ambiguous – see Fisher et al, 2006) and can therefore be assessed using, crisp probabilistic measures to give measures of error.

Two examples illustrate these problems with independence in the mapping land cover. First, any time-series of satellite imagery will contain a mixture of correlated and non-correlated information, which cannot be treated as independent (though they are often treated as conditionally independent). Second, consider how plant presence and plant cover are modelled in sample stands. It is often acceptable to consider the presence of plant species in a large sample stand of several square metres to be independent. However when the size of the sample is reduced, the presence of plants becomes positively correlated at a scale that picks out habitat patches (e.g. blanket bog with pools and dry areas). At a smaller scale still, the same species might be negatively correlated because they start to exclude one another.

4. METHODS

4.1 Formalisms

4.1.1 Bayes and Dempster-Shafer

Bayes' theorem computes the probability of an hypothesis or event, *h* given the evidence, *e* in support of that event, $P(h|e)$:

$$P(h \mid e) = \frac{P(h) * P(e \mid h)}{P(e)} \tag{1}$$

Dempster-Shafer can be considered as an extension to Bayesian statistics which contains an explicit description of uncertainty, plausibility. It assigns a numerical measure of the weight of evidence (mass assignment, *m*) to *sets of hypotheses* as well as individual hypotheses. It does not consider 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:

$$m''(C) = \sum_{\substack{i,j \\ A_i \cap B_j = C}} m(A_i) m'(B_j) \tag{2}$$

4.1.2 Bayes vs. Dempster-Shafer

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 ?'' It has at its crux the notion that the evidence can be used to vary the prior probabilities, P(*h*) and evidence either supports or refutes the hypothesis. In principal this approach can be applied to any problem involving uncertainty, assuming that precise probabilities can be assessed for all events. But, this is rarely the case. Dempster-Shafer accommodates explicit representations of uncertainty, plausibility, which equates to belief plus uncertainty. Therefore a weak belief in a hypothesis does not imply a strong belief in its negation. One of the weaknesses of Dempster's rule is that it can favour a class which has low mass in two data sets over any class that has a high mass in only one data set. The classic example is that of the two doctors, one of which is 90% certain the patient has disease A and 10% disease B; the other 90% convinced over disease C and 10% disease B. DS will give 100% support for disease B, even though neither doctor thought it likely (although this can be overcome by the use of alternative fusion rules). The point being that, it may be problematic interpreting the outcomes of Dempster-Shafer relative to evidence and the hypotheses. Description of the arguments and counter-arguments put forward by both sides of the Bayes/Dempster–Shafer dichotomy can be found in a text edited by the main protagonists from either side, Shafer and Pearl (1990) and Parsons (1994) provides a good introduction to Dempster-Shafer.

5. RESULTS

The objective was to identify the potential extent of bog Priority and Annex I habitats within Upland Heathland Phase I habitats, using some additional information and the existing Phase I survey. That is, to identify of the potential extent of Bog habitats at higher and lower grains than Phase I. The 2 analyses are:
- *To determine whether any given patch of Upland Heathland is one of the Annex 1 Blanket Bogs (7130);*
- *To identify the extent of Upland Heathland (priority habitat) that can legitimately be burned, i.e. is not Bog.*

5.1 Extent of Bog Annex I habitats

In the first example, identifying Annex I Blanket Bog habitats, there are different pieces of evidence that support a number of competing hypotheses. The evidence is the presence of NCV classes M15 (*Scirpus cespitosus - Erica tetralix*) and M16 (*E. tetralix - Sphagnum compactum*). M15 is characteristic of Annex 1 habitats Active Raised bogs (7110) and Blanket bogs (7130); M16 of Northern Atlantic wet heaths (4010) and European dry heaths (4030). Other information relating to the Phase I habitat present, soil wetness and peat depth was used to identify the likely Annex 1 habitat based on the additional evidence shown in Table 1.

From the various evidence beliefs were generated for different sets of hypotheses using Dempster-Shafer (Table 2). From the same data the Bayesian probability of singleton hypotheses were calculated (Table 3) through the combined probability that each hypothesis will pass each evidence 'test'.

Evidence	Hypotheses (Annex 1 habitats)	Unc

	4010 H1	**4030 H2**	**7110 H3**	**7130 H4**	
Heathland	0.167	0.167	0.167		0.5
Peat depth	0.233		0.233	0.233	0.3
Dry soil	0.25	0.25			0.5
Acid soil			0.25	0.25	0.5

Table 1. Evidence in support of hypotheses (H)

Bayes and Dempster-Shafer provide different answers to the question of whether this patch of land is Blanket Bog (Annex I habitat, 7130). The Dempster-Shafer results have two characteristics. First, the evidence is combined over sets of hypotheses, and second it generates an upper bound of belief (Plausibility) from the uncertainty inherent in the evidence. The results of applying Dempster-Shafer belief functions to the problem show that the set {H1, H2} has the most support, but when plausibility is considered the set {H1} has most supporting evidence. The Bayesian approach only generates support singleton hypotheses and indicates support for {H4}.

Hypotheses	**Belief**	**Plausibility**
H1	0.132	0.698
H1, H2	0.377	0.566
H3	0.057	0.170
H4	0.132	0.321
H3, H4	0.057	0.057
H1, H2, H3	0.057	0.057
H1, H2, H4	0.132	0.132
Theta	0.057	0.057

Table 2. The belief in hypotheses from Dempster-Shafer

Hypotheses	**Belief**
H1	0.062
H2	0.265
H3	0.062
H4	0.372
H5	0.239

Table 3. The belief in hypotheses from Bayes

Dempster-Shafer combines evidence over a range of hypotheses and does not allocate any remaining support (i.e. the uncertainty) to ¬Belief as in Bayes. Rather uncertainty allocated to all hypotheses or the "frame of discernment", Theta. Dempster-Shafer shows how the various pieces of evidence support different sets of hypotheses. Bayes by contrast partitions the evidence between Belief and ¬Belief. The hypotheses with only 2 pieces of evidence are the most supported, as none of the evidence supports any one hypothesis with a belief of more than 0.5 (therefore in this context more evidence equates to lower belief).

5.2 Extent of Bog Priority Habitat

In the second example, identifying Priority Habitats, some remote sensing information indicates that some landscape object (e.g. a parcel or a pixel) is bog. However, there are uncertainties associated with remote sensing information. Ancillary data is used to support the allocation of the object into a particular priority habitat class. Upland Heathland is a priority habitat and has a *one to many* relationship with the following Phase I habitats:

- Dry acid heath
- Wet heath
- Dry heath / acid grassland mosaic
- Wet heath / marshy grassland mosaic

The lower bound of the priority habitat that can be legitimately burned is given by the extent of the union of these single feature (i.e. non-mosaic) Phase I parcels. If a suspected area of burning fell within this area then there is confidence that any burnt area is not on one of the ecologically important Blanket Bog vegetation communities. If the suspected area fell within the upper bound of the Upland Heathland priority habitat then more evidence is needed to determine the belief in legitimacy.

The object is to calculate overall belief in Bog and in Heath priority habitats hypotheses using evidence weighted using ecological knowledge, in order to determine whether any burning is legitimate or not. Note, that in this case disbelief in Bog equates to belief in Heath. Each outcome is initially believed to be equally likely:
P({bog}) = P({heath}) = P({not_sure}) = 1/3

Remote sensing information indicates a 90% probability of bog, 10% Heath and a 30% something else. This could be based on field validation and the probabilities do not have to sum to unity and will be normalised. The three worlds possible must be considered in light of the remote sensing evidence:
P({bog}, $pass_{rs}$) = 0.9/3 = 0.3 (0.692) normalised
P({heath}, $pass_{rs}$) = 0.1/3 = 0.033 (0.077)
P({not_sure}, $pass_{rs}$) = 0.3/3 = 0.1 (0.231)

In this example the area of suspected burning has the following hypothetical characteristics as evidence (Table 4):
- Within the upper bound of Upland Heathland priority habitat;
- Within the upper bound of Blanket Bog priority habitat;
- It is within a conservation area (e.g. SSSI);
- Most of the area is not on steep slopes (i.e. < 25º);
- Most of the suspected area is above the treeline;
- NVC / Phase II survey data for the area indicates that the suspected area contains areas of grass, heather and mire communities (M15, M16).

Evidence	Belief (in Bog)	Belief (in Heath)	Uncertainty
Remote sensing and priors	0.692	0.077	0.231
1. Within Upland Heath Mosaic	0.1	0.5	0.4
2. Within Blanket Bog mosaic	0.5	0.1	0.4
3. Not on wet soil	0.1	0.4	0.5
4. Not on slopes	0.25	0.1	0.65
5. Below 600m	0.25	0.1	0.65
6. NVC classes M15 / M16	0.5	0.1	0.4

Table 4. Evidence supporting Bog and Heath, with ecological weighting

The normalising factor is used to update the conditional probabilities of the three classes using Bayes theorem applied to the evidence from the 6 'tests' in Table 4 using Equation 1.
P({bog} = P($pass_{rs}$).($pass_1$, $pass_2$, $pass_3$, $pass_4$, $pass_5$, $pass_6$)
= 0.692 x (0.1 x 0.5 x 0.1 x 0.25 x 0.25 x 0.5) = 0.00010817
P({heath}) = 0.077 x (0.5 x 0.1 x 0.4 x 0.1 x 0.1 x 0.1) = 0.00001538
P({not_sure})= 0.231 x (0.4 x 0.4 x 0.5 x 0.65 x 0.65 x 0.4) = 0.00312000

These are normalised by the total probability of all worlds, given all pass (0.00324356)
P({bog}|{all pass}) = 0.033493108
P({heath}|{all pass}) = 0.000476346
P({not_sure}|{all pass}) = 0.966030546

In Dempster-Shafer each piece of evidence may be combined to determine belief in bog (*Bel*), disbelief in bog, which equates to belief in heath (*Dis*) and uncertainty (*Unc*), according to the formulation from Tangestani and Moore (2002):

$$\text{Belief} = (Bel_1.Bel_2 + Unc_1.Bel_2 + Unc_2.Bel_1)/\beta \quad (3)$$

$$\beta = (1 - Bel_1.Dis_2 - Bel_2.Dis_1) \quad (4)$$

Applying Dempster-Shafer and Bayesian approaches to combine the evidence generates different overall weightings:

	Dempster-Shafer	Bayes
Belief(Bog)	0.836	0.033
Belief(Heath)	0.149	0.000
Uncertainty	0.015	0.966

These two approaches for combining information and evidence generate very different results in this instance: Dempster-Shafer has partitioned the uncertainty in the evidence into belief in bog and belief in heath (i.e. disbelief in bog) under the assumption of conjunctive evidence. The Bayesian approach assumes independence of evidence, using a multinomial probability approach effectively calculates the probabilities of a set of events that are believed to be possible, based on passing a series of tests.

Using data relating to soil wetness, elevation, slope and the presence of NVC classes, maps of different degrees of legitimacy can be constructed as in Figure 2 using the 2 approaches to combining evidence, Dempster-Shafer and Bayesian inference, in the assessment of a single hypothesis. These maps in this figure were generated without using any remote sensing information (i.e. they combine evidence 1-6).

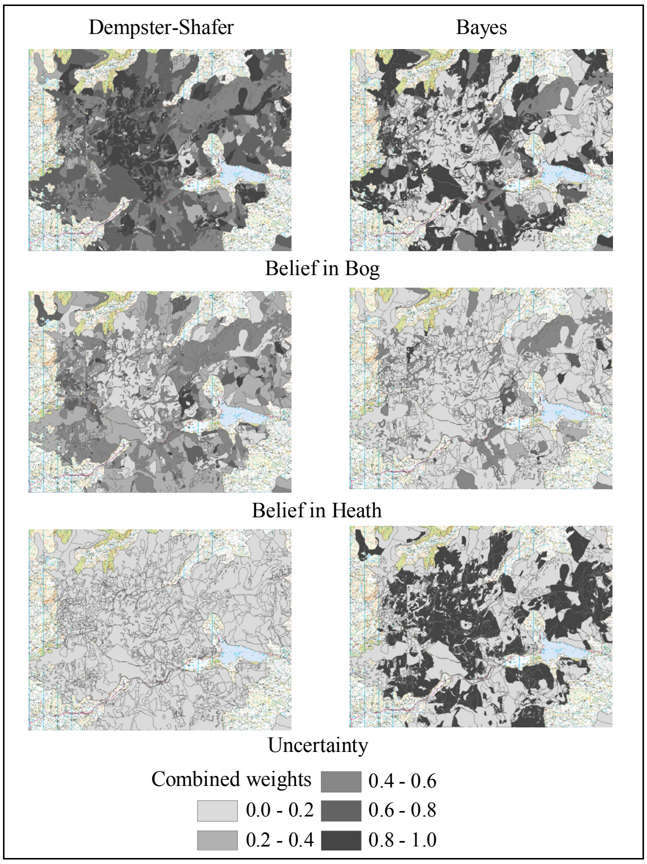

Figure 2. Upper bounds of Bog (illegitimate burning areas) and Heath (legitimate burning areas) in a test area using Dempster-Shafer and Bayesian probability, context from OS 1:25 000 Raster scanned maps (© Crown Copyright, Ordnance Survey, an EDINA Digimap/JISC supplied service)

6. DISCUSSION AND CONCLUSION

Remote sensing of land cover is an inherently uncertain exercise due to the spectral and spatial limitations of remotely sensed imagery. Because of this a number of applications incorporate other information into the classification process, including ancillary data (soils, geology, elevation) and rules relating to plant and vegetation phenological cycles. The worked examples highlight a number of issues relating to fusing different information:

1) The method by which information and evidence are combined results in different mapped and modelled outcomes;

2) Different landscape questions require different weightings strategies;

3) The method by which data at different scales and grains are combined needs to be considered in relation to weights.

This indicates that decisions relating to what features to display or map depend on the intended use of that information. Decision theory is implicit for the creation of any such map through the concept of 'expected value', where the value of a decision taken on the basis of the mapped data relates probability of each possible outcome and its value. Decision making can inform such decision making (Choquet, 1953; Chu and Halpern (2003a,b) where the evaluation of different decisions are ordinal (i.e. not cardinal), as is the case when mapping decisions are taken by conservationists and ecologists, who evaluate different elements of the landscape qualitatively rather than quantitatively: "habitat A has a greater priority than habitat B" rather than "habitat A has 3 times the value of habitat B".

Conservation agencies would like to be able to develop and produce alternative, context dependent maps relating to a series of different questions and granularities. Managers need to be able to identify the 'best' decision in the face of uncertainty. Different methods for handling uncertainty in information integration activities such as Bayesian probability and Dempster-Shafer produce different outcomes and there is also uncertainty due to the scale or grain of the object, process or question. This situation describes an interdependence between grain, scale, uncertainty and integration method. Management decisions have to be made in light of the set of possible outcomes, indicating the need for a formal evaluation of the expected value of that decision using decision theory.

7. REFERENCES

Choquet, G, 1953. Theory of capacities. Annales de l'institut Fourier, 5, p. 131-295.

Chu, F.C., Halpern, J.Y., 2003. Great Expectations. Part I: On the Customizability of Generalized Expected Utility, *Proc. 18th International Joint Conference on AI (IJCAI)*, pp. 291-296

Chu, F.C., Halpern, J.Y., 2003. Great Expectations. Part II: Generalized Expected Utility as a Universal Decision Rule. *Proc. 18th International Joint Conference on AI (IJCAI)*, pp. 297-302

Comber, A., Wadsworth, R., Fisher, P, 2006. *Reasoning methods for handling uncertain information in land cover mapping*. In *Fundamentals of Spatial Data Quality*, edited by R.Devillers and R.Jeansoulin, ISTE, London. pp123-139.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., 2005a. You know what land cover is but does anyone else?…an investigation into semantic and ontological confusion. *International Journal of Remote Sensing*, 26 (1): 223-228.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., 2005b. What is land cover? *Environment and Planning B: Planning and Design*, 32:199-209.

Comber, A.J., Law, A.N.R., Lishman, J.R., 2004a. A comparison of Bayes', Dempster-Shafter and endorsement theories for managing knowledge uncertainty in the context of land cover monitoring, *Computers, Environment and Urban Systems*, 28(4): 311-327.

Comber, A.J., Law, A.N.R., Lishman, J.R., 2004b. Application of knowledge for automated land cover change monitoring. *International Journal of Remote Sensing*, 25(16): 3177-3192.

Commission of the European Communities, 1992. Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. *Official Journal of the European Communities*, L206.

Fisher, P.F.,1997. The pixel: a snare and a delusion. *International Journal of Remote Sensing* 18 (3), 679-685.

Fisher, P., Comber, A., Wadsworth, R., 2006. Approaches to uncertainty in spatial data. In *Fundamentals of Spatial Data Quality*, edited by R.Devillers and R.Jeansoulin, ISTE, London. pp43-59

Friedl M A, McGwire K C, McIver D K, 2001, An overview of uncertainty in optical remotely sensed data for ecological applications. In *Spatial Uncertainty in Ecology* Eds C T Hunsaker, M F Goodchild, M A Friedl, T J Case (Springer, New York) pp 258-283.

JNCC, 2003. *Handbook for Phase 1 habitat survey - a technique for environmental audit, field manual*. JNCC, Peterborough, UK, p62.

Parsons, S. 1994. Some qualitative approaches to applying the Dempster–Shafer theory. *Information and Decision Technologies*, 19(4), 321–337.

Rodwell, J.S. 2006. *National Vegetation Classification: users' handbook*. Joint Nature Conservation Committee, Peterborough.

Shafer, G., Pearl, J. 1990. *Readings in uncertain reasoning*. San Mateo: Morgan Kaufmann.

Tangestani M.H., Moore F. 2002. The use of Dempster-Shafer model and GIS in integration of geoscientific data for porphyry copper potential mapping, north of Shahr-e-Babak, Iran. *International Journal of Applied Earth Observation and Geoinformation*, 4: 65-74.

UK Government, 1994. *Biodiversity: the UK Action Plan*, HMSO, London.